Features of Dialogic Instruction in Upper Elementary Classrooms and their Relationships to Student Reading Comprehension

Author: Catherine Michener

Boston College
Lynch School of Education

Department of
Teacher Education

Curriculum & Instruction

FEATURES OF DIALOGIC INSTRUCTION IN UPPER ELEMENTARY CLASSROOMS AND THEIR RELATIONSHIPS TO STUDENT READING COMPREHENSION

Dissertation
by
CATHERINE J. MICHENER

submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

August, 2014

This dissertation was presented

by

Catherine J. Michener

It was defended on

April 9, 2014

and approved by

Dissertation Advisor: C. Patrick Proctor, EdD., Lynch School of Education, Boston College

Maria Estela Brisk, PhD., Lynch School of Education, Boston College

Patrick McQuillan, PhD., Lynch School of Education, Boston College

Gina Biancarosa, EdD., College of Education, University of Oregon

Abstract

There is widespread agreement that language skill underpins reading comprehension (e.g. Cutting & Scarborough, 2006; Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Ouellette, 2006; Snow, 1991), and empirical work over the last 20 years has shown positive effects of dialogic instruction on student literacy outcomes. This suggests the importance of the engagement with others in the learning process as a scaffold for academic literacy skills (Wells, 1999). Research in this area has shown a number of important features of dialogic instruction to be positively correlated with literacy skills, but it is still not well understood how teachers guide and support students in developing language abilities for reading comprehension. Drawing on dialogic theories of language and the simple view of reading model (Hoover & Gough, 1990), this study pulls together mechanisms of productive discourse from the literature to test their efficacy in literacy instruction.

Dialogic instruction has been researched through intervention studies, large-scale correlational studies, and observational studies using discourse analysis. In the current study, the latter two traditions are drawn upon to explore the nature of dialogic instruction in 31 classrooms across grades three, four, and five. Using a convergent mixed method analysis, the study explores how features of dialogic instruction relate to students' reading comprehension outcomes, and identifies themes within the patterns and variations of these features during instruction. Multilevel modeling (Raudenbush & Bryk, 2002) and case study analysis (Merriam, 1998; Stake, 2006; Yin, 2009), including the identification of patterns and themes (Coffey & Atkinson, 1996), are used to identify significant talk moves for reading comprehension and to qualify the content and function of these moves in their instructional contexts.

Quantitative analyses showed five significant talk moves predicted reading comprehension achievement, including the rate of uptake questions, teacher explanations, and low-quality evaluations. High rates of student explanations and high-quality questions were predictive of lower reading outcomes. Case study analyses of each finding suggest some common patterns and themes in these language arts lessons, including a preponderance of teacher talk, a lack of quantity and quality to student talk, and an efferent stance (Rosenblatt, 1994) toward reading. These findings indicate a lack of dialogic practices across the grades and classrooms. However, within the pervasive IRE instructional discourse pattern (Cazden, 2001; Mehan, 1979) of these lessons, there were opportunities for dialogic practices that support

students' linguistic comprehension. Uptake questions functioned to support student linguistic development by focusing student attention on inferencing skill while creating opportunities, albeit limited, for language output. High rates of talk moves that function as sources of academic language exposure, including direct teaching moves like teacher explanations, explicate some ways teachers can support student linguistic comprehension, a key component of the simple view of reading. Overall, this analysis showed mixed results for the importance of dialogic instructional moves, and indicates the importance of teacher talk to facilitate linguistic comprehension, as well as the promise of talk moves that incorporate student attention and participation around texts.

Acknowledgements

This research study rests on the assumption that learning is fundamentally a collaborative activity, and this study is no exception. Despite some agonizing hours and weeks and months alone at my desk, this work was built and accomplished because of the collaboration and support from many people.

I am so privileged and grateful to have the support of an incredible committee of scholars and mentors. My inexpressible thanks go to Patrick Proctor for his unwavering support, patience, and challenging ideas. I consider myself so fortunate to have Patrick as my mentor; it has been a privilege to work with him. Patrick's mentorship has taught me how to write better, work better, and *be* better. My work with Maria Estéla Brisk has encouraged me to keep focused on the classroom, and her incisive feedback and patience gave me the support and confidence to do this work. I am grateful to Patrick McQuillan for building my understanding of the importance of qualitative methodology to answer important questions, and for guiding me to conceptualize the larger scope of this work. And finally, I wish to express my sincerest appreciation to Gina Biancarosa for her seemingly infinite patience with these analyses, and for providing me with detailed feedback and her incisive analytic thinking about these findings.

This study would not have been possible without the CLAVES project. I am indebted to Patrick Proctor and Rebecca Silverman for their support and trust in using CLAVES data to further explore the teaching and learning in these school sites. To the Boston College CLAVES research team, I had so much fun working and learning with you. Thank you for all the discussions about teaching, schooling, and feasible research projects while we drove back and forth in the rain, sleet, and snow! I also wish to thank the generosity of Anna Meyer for introducing me to her school sites. My sincerest appreciation goes to the teachers and students who so kindly shared their classrooms and their time with me to help the project collect surprisingly complete data on which this research draws. Being a part of their classrooms and school communities was a true privilege.

And finally, I wish to thank my cohort for helping me in innumerable ways to get through this program successfully, and to my dear friends in previous and following cohorts who continue to challenge me to be a better person and researcher with their own research and successes. To my friends Julie, Olga, Nelly, Rebekah, Sarah, Diana, Susan, and Erin, I thank you

for sharing your strong selves to provide me with guidance, support, and a sense of humor about it all. I also thank Josh for helping me get started on this project, my family for pushing me through to the end, and my sincerest appreciation to Quanah for your patience and wit. I hope you jump at all the opportunities that come your way, and continue to challenge yourself to do good in this world.

Table of Contents

List of Tables

List of Figures

Chapter 1 – Overview of the Study

> *Language remains the most fundamental resource in which participants negotiate and construct their meanings in classrooms* (Christie, 2002, p. 10)

Introduction and Rationale

The enterprise of education is the intersection between curricula, instruction, and learning, and this is accomplished fundamentally through language. We teach through language, learn through language, are assessed through language, and we research this teaching and learning through language. Language permeates teaching and learning, and it is no wonder that the language of instruction is a continued focus of educational research. One of the most fundamental points about the language of instruction is that it structured experience (Christie, 2002), routinized with time and the expectations of its participants. This structured experience is constructed by the participants, the teacher and the students (Christie, 2002). Therefore, there is a paradox to the language of instruction and the goals of learning in the classroom: the study of language in classrooms has a focus on inter-individual talk, and yet the goal of schooling is individual learning (Cazden, 1988). Thus, to study the language of instruction is to also consider "how observable classroom discourse affects the unobservable thought processes of each of the participants, and thereby the nature of what all students learn" (Cazden, 1988, p. 99). This intersection between language and learning outcomes is a complicated but important place to explore if we are to ameliorate learning outcomes for all students.

One current area of attention of those interested in improving the learning opportunities of students in the US is literacy development. This has been a heated topic, exemplified by the

controversies over the National Reading Panel report (NIH, 2000) and tensions over the later report of the National Literacy Panel on language-minority youth (August & Shanahan, 2006). No matter the political nature that literacy instruction has taken however, literacy instruction is still a vital area in which to research, as reading is almost as ubiquitous in learning as language in K-12 classrooms. Indeed, by middle elementary school, academic language and literacy skills are "presupposed by and involved in almost every kind of school learning," and that poor academic achievement in middle and high school may well be rooted in elementary years (Snow, Barnes, Chandler, Goodman, & Hemphill, 1991). The features of instruction for successful reading achievement are undoubtedly varied and context-dependent, and many in use are not empirically supported. This is due, in part, to our developed but still incomplete understanding of the reading process, as indicated by the various models of reading that continue to be created, revised, and critiqued. Research exploring reading instruction is thus one area that continues to be of import, not only because it could inform theories of the reading process, but also its implications for the many students struggling to understand texts demanded by current curricula.

Currently, there is an interest in the nature of instructional discourse as one area of promising pedagogical practice for successful reading achievement. Instructional talk as a research focus has been said to be a "rich lens for analyzing….teacher-student interaction in teaching and learning" (Wolf, Crosson, & Resnick, 2004), and research on literacy learning has heeded this perspective. The link between language and literacy learning is not surprising, since language and literacy research has established that early language skills are key precursors to literacy development and achievement (Dickinson & Tabors, 2001). Though we know that a variety of factors, such as socioeconomic status (e.g.Walker, Greenwood, Hart, & Carta, 1994) and vocabulary (e.g. Hart & Risley, 1995), show strong correlations with reading achievement,

more recent work has indicated that the language in which early readers are immersed also predicts literacy achievement (Dickinson & Tabors, 2001). The classroom is thus one important source of language immersion and its potential effect on the development of reading skills pushes the educational community to better understand the role of language in the development of children's literacy.

In particular, dialogic instruction, informed by sociocognitive and sociocultural theories, has been one feature of instructional talk linked to gains in student reading comprehension. Dialogic instruction is a nebulous concept not well defined in the literature and not well enacted in typical classroom practice, as many researchers have observed (e.g. Cazden, 1988; Mehan, 1979; Nystrand, Gamoran, Kachur, & Prendergast, 1997), but there have been promising findings as to its effects on literacy outcomes, and interesting suggestions for why dialogue may support students' work with texts. Before we delve into this research however, more must be said about the literacy contexts in which dialogic instruction may act as an important scaffold.

National data in the US continues to show large achievement gaps between students' reading abilities. For example, on the National Assessment of Educational Progress (NAEP) reading test, only 4% of English language learners (ELL) in the nation scored at the proficient level in eighth grade, while 31% of non-ELL students scored at the proficient level (Lee, Grigg, & Donahue, 2007). Other groups of students exhibit similar achievement gaps. In addition to standardized test scores, curricular transitions also reveal group disparities in reading achievement. The fourth-grade reading gap, as it has come to be known, points to the realities of students struggling to meet a changing curriculum and a changing focus of reading instruction. Decoding and building fluency are foundational reading processes learned in the early stages of literacy acquisition, but by fourth grade, students are expected to have become independent and

strategic readers and use their reading skills to engage with text in sophisticated ways; students are expected to analyze, synthesize, organize and evaluate a wide range of opinions and ideas presented for texts in various media (National Council of Teachers of English, 2006b). As students progress through school, the success of reading relies heavily on the knowledge of decontextualized vocabulary and the language of the text, and less on context (Snow et al., 1991). By the middle elementary grades, curricula demand text-level skills that require readers to glean significant amounts of conceptual knowledge from reading increasingly abstract texts (Lesaux & Geva, 2006). This change in reading goals, from learning to read to reading to learn (Chall, 1983), demands changes in reading instruction to mitigate the challenge posed by this transition for many students.

An additional challenge for instruction is the shifting nature of the demographic character of US public K-12 schools. One in five children speak a language other than English at home (KewalRamani, Gilbertson, Fox, & Provasnik, 2007), and 10% of the total K-12 student population is classified as ELL (NCELA, 2008; NCES, 2013), students identified as needing language support to complete grade-level content in English. Elementary educators are thus faced with increased public scrutiny on their students' reading outcomes, a curriculum that shifts goals though many students may not yet be ready to transition along with it, and increased linguistic and cultural diversity to which many educators are unprepared (Darling-Hammond & Bransford, 2005). It is for this confluence of factors that research into effective reading instruction is critical if educational research is to serve teachers and students well. Along with a lack of attention to reading instruction at the upper elementary grades, this situation suggests a striking need for research in this area. This is the expressed interest of the current proposed study.

Theoretical and Conceptual Framework

The assumption undergirding this study is that interaction through language supports reading comprehension. This is a foundational claim of sociocultural theory in education: the process of learning, understanding, and knowledge-building are inherently social activities that typically occur in culturally- and instructionally-mediated contexts (Bransford, Brown, & Cockling, 2000; Cazden, 1988; Palincsar, 1998; Sawyer, 2006; Vygotsky 1978).

Greeno (2006) has argued that all "socially-organized activities provide opportunities for learning to occur" (p. 80). From a linguistic perspective, sociocultural theory rests on the importance of language in the learning process: "in the development of the child as a social being, language has a central role. Language is the main channel through which the patterns of living are transmitted to him, through which he learns to act as a member of a 'society'… and to adopt its 'culture', its modes of thought and action, its beliefs and its values" (Halliday, 1993, p. 9). Sociocultural theory enacted in empirical research in school settings suggests that central to social interaction in the classroom learning environment is the collective discourse in which students and teachers engage. Cognitive development is seen as a process of meaning making with others (Halliday, 1993; Vygotsky 1978). Sawyer (2006) has argued that interaction is "the place where group knowledge translates into individual cognitive advancement" (p. 190). Others have suggested that interacting with a group creates something more than what the individuals in the group possess in themselves (Forman & Cazden, 1994; Palincsar, 1998; Sawyer, 2006; Soter et al., 2008). Similarly, Scardamalia & Bereiter's (2006) body of work reflects an attention to the individual learning process by focusing on how individuals contribute to group knowledge, arguing that learning is first external and collective, and then becomes internalized. Others have

suggested that knowledge is never truly internalized because its assessment and function exist in a context of action and sociocultural interaction (Lave & Wenger, 1991; Rogoff, 1995).

Learning theory from a sociocultural standpoint argues that formal and informal learning is manifest through interactions mediated by cultural norms, and how this learning actually occurs rests in the realm of sociocognitive theory. Sociocultural and sociocognitive theory are very similar frames of reference, so much so that some scholars have treated them as synonymous in their work (e.g. Applebee, Langer, Nystrand, & Gamoran, 2003). For instance, Murphy and colleagues (2009) have presented literacy research that focused on text-based discussions to support comprehension, situating these studies sociocultural tradition. The authors of a pedagogical framework emphasizing discussion drew from a sociocognitive perspective on dialogic instruction, explaining that it is the view that individuals bring attributes to the group, but the group is more than the sum of the abilities of the individuals: the individuals in a group form an interpretive community, a context that influences the thinking and learning of each individual (Chinn, Anderson, & Waggoner, 2001). In the Dickinson & Tabors (2001) study, the researchers were explicit about exploring cognitive processes in a social context, looking at language that "replicates some of the demands of literacy- that is, talk that requires participants to develop understanding beyond the here and now and that requires the use of several utterances or turns to build a linguistic structure, such as explanations, narratives, or pretend" (p. 2). The work on discursive practices developing reading skills thus draws from both traditions, perhaps because the processes of reading have been most persuasively explained through models that emphasize cognitive processing, whereas studies of instructional discourse have emphasized the social and cultural contexts of classroom talk. Gee (2001) has argued that these theoretical

viewpoints have pushed notions of literacy from a narrowly defined psycholinguistic processing skill toward literacy as language that demands specific verbal abilities in different contexts.

Both sociocultural and sociocognitive perspectives emphasize the benefits of various forms of discourse to be used in schools; the sociocognitive perspective focusing on understanding complex academic content and the opportunities of access to this content for all students (Michaels, O'Connor, & Resnick, 2008), and the sociocultural on the social interaction and discourse in which classroom participants engage. Therefore, research on classroom discussion has been studied from perspectives that suggest that social actions and discourse play fundamental roles in cognitive development (Vygotsky, 1986). Another way of describing the importance of understanding classroom talk is to say that, fundamentally, speech unites the cognitive and the social (Barnes, 1974, as cited in Cazden, 1988).

To better understand the contexts of this literacy instruction, the current study of classroom discourse is rooted in sociocultural theory, drawing on these theoretical frameworks like other researchers interested in discourse as learning (Chang-Wells & Wells, 1993; Gee & Green, 1998; Palincsar, 1998). The current study explores the nature of dialogic instruction in language arts contexts through quantitative and qualitative analyses. By focusing on teachers' dialogic practices, a sociocultural perspective emphasizes how instructional discourse functions: "a pedagogical discourse operates by taking forms of knowledge from elsewhere and 'relocating' these for the purposes of the initiation of others" (Christie, 2002, p. 24).

In order to explore the nature of effective instruction as it relates to student achievement through a sociocultural lens, a primary focus must be on classroom discourse. The terms used in various studies of classroom talk have drawn on both explicitly and vaguely defined terms, so some clarifications are needed. The following are the key terms used in the proposed study.

Key terms.

Discourse has been defined as "language in time," shaped and structured by interlocuters who respond to each other reciprocally (Nystrand, Wu, Gamoran, Zeiser, & Long, 2003). Classroom discourse has been described as "all forms of language, including gesture, signs, artifacts, mimicking" and provides "information about articulated or unarticulated teacher expectations for student participation and learning" (Lerman, 2001, p. 88). Cazden described classroom discourse as "situated language use in one social setting" (Cazden, 1988, p. 3). Thus, discourse is a general term used to define a number of communication acts usually in a particular setting. Intentionally broad, this definition needs more specification to be operationalized in a research study with limits. A more operational definition, *oral discourse* has been defined as "extended oral productions, whether monologic or multi-party, centered around a topic, activity, or goal" (Lawrence & Snow, 2011, p. 323). With this definition of oral discourse, one can then begin to consider, through either a sociocultural or sociocognitive frame, how it relates to literacy. For instance, are their specific forms of discourse that relate to the reading process, or do general discourse skills draw on language skills in support of reading (Lawrence & Snow, 2011)?

Discussion has been a commonly used term used in empirical studies exploring this phenomenon of classroom discourse. It has been described in dichotomous terms with recitation, a pattern of classroom talk characterized by Mehan (1979) and Cazden (1988) as initiation-response-evaluation (IRE). Discussion, on the other hand, is "when classroom discourse transcends the typical question/response/evaluation sequence to exhibit free-flowing comments among students and the teacher" (Gamoran & Nystrand, 1991, p. 281), or more tightly defined as the "free exchange of information among students or between at least three students and the

teacher that lasted longer than 30 seconds" (Nystrand & Gamoran, 1991, p. 274; Nystrand et al., 1997). Dillon (1994) provides a thorough overview of the concept of discussion generally and specifically related to classrooms. The author notes that the term discussion "covers an ill-defined range of talk" and is used colloquially as just about any kind of talk, from conversation, recitation in classrooms, debates, and lecture (p. 6). He defines discussion at length:

> Discussion is a particular form of group interaction where members join together in addressing a question of common concern, exchanging and examining different views to form the answer, enhancing their knowledge or understanding, their appreciation or judgment, their decision, resolution or action over the matter at issue (p. 8); a form of disciplined and concerted talk (p. 13).

Dillon notes that the importance of discussion drawn upon both classical intellectual inquiry and liberal democratic theory (see also Michaels et al., 2008), and follows principles like logic, reasonableness, diversity of perspectives, and open time limits. Wolf et al. (2004) write that rigorous discussion "reinforces students' understanding of a challenging text or concept" (p. 30). Dillon juxtaposes this with the common classroom discourse pattern of recitation, where the teacher is the predominant speaker and there is little in the way of exchanging and examining differing views. Discussions typically have few questions and most questions are clarifications and authentic, and display regular uptake, where previous utterances are connected to current ones, either by teachers or students building on each others' comments (Nystrand & Gamoran, 1991). Discussion in classroom context encapsulates the idea of reciprocity, where interlocuters incorporate each others' perspectives in interaction.

Dialogic instruction is generally defined in opposition to monologic instruction or a transmission model of teaching (Wells & Arauz, 2006). It is reciprocal in nature, like the term discussion has been described, but dialogic instruction has been used in the literature as a broader term than operationalized definition of discussion. This instruction is characterized in the

literature as exhibiting features like those used to define discussion, but is generally treated more as a pattern of discourse that is supportive of pedagogical activities like text-based discussions and inquiry work (e.g. Applebee et al., 2003). Features of dialogic instruction that have been empirically studied are pedagogical talk moves such as the nature of teacher and student questions, teachers evaluations of students responses, the reference of previous ideas to current ones, and the ratio of student to teacher talk.

In this study, we hypothesize that features of dialogic instruction support r*eading comprehension.* There are many models of reading comprehension, which will be discussed in more detail in Chapter 2, but it is important to note here that there have been many variables proposed key to the comprehension of texts: evidence of prior knowledge, cultural background, genre knowledge, motivation, vocabulary, linguistic comprehension, among others. The model of reading comprehension used in the current study is the simple view of reading (Hoover & Gough, 1990), in which decoding and linguistic comprehension are proposed as the two fundamental elements for reading comprehension. How linguistic comprehension can be facilitated through instructional talk is the focus of this study. The implications of effective reading comprehension has been characterized as allowing students to "participate in disciplinary conversations" to do well in a range of school tasks (Applebee et al., 2003, p. 695), a distinctly discursive description of reading.

A final concept that figures in this study is the nature of dialogic instruction in multicultural and multilingual classrooms. As public school classrooms are characterized by increasing diversity, reading instruction must adapt to the needs of students who have different backgrounds and resources than monolingual, culturally-dominant students. For dialogic instruction to support the development of reading skills, educators must have some

understanding of second language acquisition for oral language and literacy development. Language-minority students can be more dependent on teachers and their peers for access to cultural and linguistically appropriate forms of language and behaviors. Academic language use, and opportunities for academic talk like text-based discussions allows for the "intellectual socialization" (O'Connor & Michaels, 1993, p. 319) of students, which is a particularly important scaffold for student who do not have this 'academic' linguistic capital from their home contexts. Thus, exploring the possible interaction of dialogic instruction and language status may be fruitful in understanding effective supports for this growing student population.

Purpose of the Study

This study is designed to describe current literacy instruction characteristics and explore the impact of dialogic instruction as a classroom-level predictor of student-level reading comprehension outcomes. Drawing on a sociocultural frame, attention to effective instructional practices for literacy necessarily means studying the various patterns of discourse exhibited in learning environments. Nystrand (2006) and others (e.g. Snow, 1991) have suggested that literacy outcomes cannot be understood through a direct line from instruction to outcome measures, and therefore it is incumbent upon the researcher to explore the complexity of possible variables alive in the instructional environment, and how these might relate to student-level learning outcomes. To do this, a mix of methods is used in the current study. As there are few empirical studies of upper elementary language arts lessons, and few that use an integration of mixed methods to better understand the complexity of classroom language and literacy contexts and their correlated outcomes, there is a need to provide qualitative descriptions of literacy instruction and quantitative measures of outcomes in a policy environment where curricula are making increasing demands on students' abilities to read. The research questions (below)

guiding this study are meant to address this gap in the literature, and to contextualize correlational relationships of talk moves and reading comprehension. To answer research question one, mechanisms of dialogic instruction are identified in instruction and used in hierarchical linear models that account for students clustered in classes. In this manner, we will be able to explore classroom-level predictors of individual-level achievement, and explore any differential impacts of dialogic instruction for bilingual and monolingual students. Research questions 2 and 2a contextualize these correlational findings by looking back into the lessons to examine their content and function more closely and qualitatively, and to provide more explanatory power of the relationship between talk moves and student reading achievement.

Research Questions

Overarching Research Question: How does the quantity and quality of dialogic instruction affect student reading comprehension in culturally- and linguistically-diverse, upper-elementary classrooms?

RQ 1: What features of dialogic instruction predict reading comprehension scores?

RQ 1b: Are bilingual and monolingual students differentially impacted by exposure to high levels of dialogic instruction?

RQ 2: What patterns and variations in the significant findings of dialogic instruction from RQ1a are seen in these culturally- and linguistically-diverse classrooms?

RQ 2a: What variations of significant features of dialogic instruction are seen within classrooms?

Chapter 2 – Review of the Literature: Instructional Discourse and Reading

The current study seeks to contribute to our understanding of effective classroom discourse for reading comprehension in upper elementary, multicultural and multilingual classrooms. By implication, this study aims to generate knowledge to support educators in literacy instruction for both monolingual and linguistically-diverse learners, as this is the new US norm in K12 classrooms (Capps, Fix, Ost, Reardon-Anderson, & Passell, 2004; KewalRamani et al., 2007). The body of literature on instructional talk and literacy outcomes is extensive, and as such, cannot be covered in a single review. The focus of this review is on reading comprehension outcomes and the evidence for features of dialogic instruction that have been empirically pursued to relate to reading and reading-related outcomes in K12 students in English-medium schooling settings.

Empirical and conceptual studies included in the following literature review were drawn from searches on ERIC, ERC, Google Scholar, JStor and PsycInfo using the following search terms: *classroom discourse, classroom talk, instructional talk, dialogic instruction, inquiry, discourse analysis, literacy & literacy achievement, literacy outcomes, reading comprehension, literacy instruction, elementary & secondary students, bilingual students, English language learners,* and *mixed method*. Excluded from this review are studies that have used critical discourse analysis. Studies using critical discourse analysis in US classrooms tend to be more focused on the positions of the learners in relation to one another. While this is undoubtedly an important influence on student learning, the studies discussed here have more explicit foci on academic, content-area achievement outcomes. Additionally, we have tried to consider studies conducted in mainstream classrooms, rather than specialty or English-as-a-Second-Language (ESL) classrooms, as the focus of the study is on dialogic instruction, using data from upper-

elementary language arts lessons. Research in the last two decades will be emphasized, with the recognition that recent research has a robust foundation in works of Bellack et al. (1966), Flanders (1970) and Cazden, John and Hymes (1972), three of the first works to explicitly study classroom discourse to understand and improve classroom teaching (Christie, 2002).

This chapter begins by reviewing the theoretical foundations of current work on classroom discourse, followed by a discussion on current theoretical underpinnings of reading comprehension, the study's outcome variable. This section is followed by a review of research on reading instruction and of the empirical work on features of effective literacy instruction. The chapter concludes with a review of mixed method studies in literacy research. This review includes five features of instructional discourse that have figured prominently in the research (question type, question quality, evaluation, explanation, and first language use). Included are findings in both elementary and secondary contexts. Studies that have explored discussion in classrooms with bilingual learners are interwoven throughout the review.

Theoretical Foundations of Research on Classroom Discourse

The importance of language exposure and language use to cognitive and academic development stems from sociocultural and socioconstructivist theories of learning that emphasize the social nature of learning through interaction with others and with the learner's environment (Herrenkohl & Mertl, 2010; Lave & Wenger, 1991; Mercer, Wegerif, & Dawes, 1999; Rogoff, 1995; Vygotsky 1978; Wells, 1999; Wells & Arauz, 2006). Drawing on these general learning theories, classroom language has been an explicit focus in literacy and general learning research over the last 30 years, as witnessed in the successive editions of *The Handbook of Reading Research*. In the latest review of empirical work on oral discourse and reading outcomes, Lawrence and Snow (2011) note that the Vygotskian claim that "oral discourse is a context for

practicing, learning, and appropriating comprehension skills, has been posited as valid across ages and levels of reading skill" (p. 321). Through the oral interaction and scaffolded oral interaction of interlocuters, children develop language and conceptual knowledge. The process stems from talk and new language and concepts are integrated into the individual's cognitive system.

Vygotsky's argues that higher mental functions first manifest in social interaction, and then are integrated into individual cognitive structures as the process for conceptual and linguistic learning. Rogoff (1995) and Lave and Wenger (1991) have extended this view of learning to blur the separation of the social and individual realms to argue that all learning is social, always happening through participation and negotiation with others. In this sense, learners do not integrate externally derived concepts into their existing cognitive structures, but appropriate a given practice such as literacy by transforming "their understanding of and responsibility for activities through their own participation'' (p. 150). In this view, teacher talk is a form of guided participation for appropriation into a community of practice like the literate work of school. Halliday's (1993) contributions to a theory of learning makes the link between language and learning perhaps more explicit: learning is inherently a meaning-making, semiotic process, so when children learn language, "they are learning the foundation of learning itself" (p. 93). In essence, language is learning.

In a typical Western classroom context, this language is a structured experience where certain language patterns, like questioning, directing, responding, and evaluating, work through language to frame learning that can constrain or provide student with opportunities to develop their linguistic and conceptual knowledge (Christie, 2002). The Initiate-Respond-Evaluation (IRE) pattern in classrooms (Cazden, 1988; Mehan, 1979) is a genre of talk that is common in

classrooms and has been critiqued for constraining student talk (Cazden, 1988), and considered for its potential to initiate more open and exploratory talk (Christie, 2002; Nassaji & Wells, 2000; Wells, 2000). Empirical work by ethnographers like Barnes and Todd (1977) and Heath (1982, 1983) have supported these theoretical claims of social construction and patterns of language in and out of school that affect the language of schooling and by extension, student learning.

With regard to literacy learning in particular, however, there is an insufficient amount of research to support the claim that oral discourse in the form of classroom discussions and like instructional discourse contributes to literacy skills (Murphy et al., 2009), particularly with few studies of unprogrammatic discourse and student reading outcomes at the upper elementary levels. Despite the extant lack of empirical support for these theoretical positions, sociocultural and constructivist perspectives are widely held supporting academic discourse and student learning.

Empirical Work on the Role of Classroom Discourse: General Findings

These theoretical positions on learning draw evidence mainly from studies of child language development (e.g. Hoff, 2006; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Snow, 1972, 1991). Of interest for the current study is the evidence of student learning outcomes from classroom talk, both with teachers and peers. Research in elementary classrooms has shown how language structures can be appropriated to scaffold learning and literacy tasks, and in some cases, how language structures can spread between students and teachers. For instance, studies of Collaborative Reasoning (CR; Waggoner, Chinn, Yi, & Anderson, 1995), a peer-to-peer, text-based discussion activity, have provided some important evidentiary support for the sociolinguistic nature of learning key academic language skills like constructing persuasive

arguments. CR as a reading program is one attempt to move away from the typical recitative format of classroom discourse in US classrooms, and emphasizes the role of students as interpretive authorities of text and managers of turn-taking, which, in recitative patterns of instruction, are features controlled by the teacher.

For example, Anderson and colleagues (2001) examined transcripts of these small-group discussions by fourth-grade students. The data indicated that when a student used a successful argument form (stratagem; e.g. "If [ACTION], then [BAD CONSEQUENCE], so [NOT ACTION]"), it tended to spread and occur more frequently. The authors suggest that students appropriated arguments stratagems when they understood them, or perhaps when the arguments were directed specifically at them. Though this conclusion was speculative, support for the importance of social interaction for this kind of cognitive and academic development also comes from a later study of CR discussions and the use of analogies. Students' use of analogies to explain or persuade peers spread during peer-led discussion groups. The study showed this rhetorical device "snowballed," or increased in frequency during discussions as new analogies were used, demonstrating students' improving understanding of the use of analogies (Lin et al., 2012, p. 1440). This "snowball" phenomenon was evident in earlier studies in CR activities (Anderson et al., 2001). These studies clear show students appropriate language forms from each other to persuade each other of their perspectives on the text. This approach to instruction and research draws on a sociocognitive perspective of learning and literacy development, emphasizing the different and divergent contributions of individuals to the group discussion, which form the social context and influences each individual to jointly construct meaning of the text (Chinn et al., 2001).

Eeds and Wells (1989) also explored features of literature discussion in a group of fifth- and sixth-grade students working with pre-service teachers, and found the highest percentages of utterances by students were text interpretation remarks and supports, while the teachers' utterances had more to do with conversation maintenance. While the discussions reflected efferent readings of text (Rosenblatt, 1994), and not aesthetic or critical interpretations (like CR discussions tend to do), Eeds and Wells showed evidence that the students and teachers built meaning together, corroborating CR findings. This meaning making supported both the teachers' learning of students' capabilities and text-based discussions and students' understandings of the text (Eeds & Wells, 1989). Unfortunately, no literacy-based outcome measures were used in the study, but the authors concluded that inclusive and rich discussions were possible with novice teachers and students with varying reading abilities.

This group of studies provides some important evidence of the social nature of upper-elementary students' language and cognitive development, particularly of argumentation, a key skill in later elementary and secondary reading and writing curricula, and in the new Common Core State Standards (National Governors Association Center for Best Practices Council of Chief State School Officers, 2010). The use of argumentation has been a focus of discourse studies in science classrooms, resulting in similar findings. For instance, one of the few intervention studies employing discourse analysis in a secondary context analyzed the role of student questions and argumentation in science content classes (Chin & Osborne, 2010). The sample was drawn from four classes of 12-14 year-olds, two classes in Singapore and two in London. Unlike the quantitative methods used in the CR studies, these researchers used a inductive, grounded approach (Glaser & Strauss, 1967) to identify patterns in the data, and coded for the depth of explanation, the elaborateness of warrants, examples used in justifications, and

rebuttals. These descriptions were linked with the quality of the questions, suggesting that not only did student questions scaffold their thinking, but student questions also elicited productive argumentation that was not evident without preceding questions.

Studies of teacher talk have also provided some evidence for the importance of language exposure for language and academic skills. Studying the social nature of talk through fine-grained discourse analysis, Maloch's (2002) qualitative case study of one third-grade teacher's classroom talk demonstrates the importance of modeling effective discussion structures beyond recitative patterns. In this classroom, the teacher's scaffolding of discussion norms showed evidence of being appropriated by students through their talk. Similarly, Sharpe (2008) looked at academic language use in classroom talk, and found that the teaching strategies of repeating a student's answer, recasting a student's contribution into more technical and academic language, and recontextualizing language by shifting to a more appropriate register for historical inquiry developed students' technical language in this academic context. Presumably this would support students' reading comprehension of history texts, but this was not an outcome measure of the study.

In other elementary and middle school classrooms, Wolf, Crosson and Resnick (2004), studied naturalistic classroom discourse using researcher-developed rubrics, the Instructional Quality Assessment (IQA), to assess whole-group pedagogical talk in ethnically-diverse classrooms grades 1-8 (N = 21 classrooms). The IQA included observational rubrics focused on accountable talk and clear expectations in the classroom, the rigor of the lesson, students' self-management of their learning, and effective teacher questioning that supported rigorous reading comprehension. Their findings described classrooms where teachers usually led the discussions, and where talk moves attended to knowledge and reasoning, but not to linking interlocutors'

ideas to one another. Talk moves had a strong positive relationship with the level of academic rigor in the lessons: when teachers and students questioned each other's knowledge, and explored one's own thoughts, these moves had positive correlations with academic rigor during reading comprehension lessons (teachers questioning: $r = .79$, student questioning: $r = .84$, and students providing rigorous thinking $r = .80$).

Wolf and colleagues' (2004) more general finding concur with intervention and other observational research that teacher talk moves are important in generating student talk, shown in the high correlations between the same kinds of talk moves between teachers and students. Unfortunately, the study did not correlate the rubrics or rigor of the reading lessons with student outcomes, which would have traced the construction of a rigorous lesson from teacher and student talk to student learning achievement to better understand some impactful variables of classroom lessons. As it was, the qualitative analyses of the discourse revealed some important talk moves related to increasing student talk, but not to an understanding of increasing reading comprehension levels. Another limitation of this study was the limited number of observations on which these rubric ratings were based. Despite prompting the teachers to demonstrate a typical lesson that involved a text read aloud, whole-group discussion, and small-group or independent work, with only one observation of a reading-comprehension lesson on which to rate lesson rigor and discourse quality, the naturalistic nature of these "typical" lessons may have been lost.

This empirical work, for both intervention and naturalistic studies of discourse, shows the value of classroom talk, both from the teacher and from peers, to scaffold student understanding, forms of rhetoric (e.g. argumentation, analogies), and academic language. While these findings indicate strong support for dialogic theories of learning, only by implication are these language

results potentially supportive of reading development. In tandem with these findings, reading comprehension has been redefined over 30 years to be conceptualized as a more dialogic process between the reader and text, drawing on the same theoretical foundations as explanations for the importance of oral discourse in cognitive and academic development. With the "dialogic turn" in research on reading comprehension (Wilkinson & Son, 2011), there has been increasingly more research and some evidence supporting the importance of social interaction for literacy achievement.

Current Theoretical Foundations of Reading Comprehension

Reading comprehension has been defined in many ways, informed by different theoretical perspectives on how comprehension of text is accomplished and applied to explain reading for both monolingual and bilingual readers. From a cognitive processing and constructivist perspective, reading comprehension is the act of "simultaneously extracting and constructing meaning through interaction and involvement with written language" mediated by the context of reading and the experiences of the reader (Snow, 2002, p. 11). Similarly, Duke and Carlisle (2011) argued that comprehension is the act of creating and adjusting a "mental representation of the meaning of the text" (p. 200). This view stems from the work of Walter Kintsch, who proposed a Construction-Integration model of text comprehension in response to schema theories that suggest a top-down driven model of reading (Kintsch, 2004). Learners construct mental representations, whereby there is some change in a learner's understanding of the world as a result of reading a text. The comprehension of text is a process of both top-down and bottom-up processing, where readers attend to word- and sentence-level features of text in their short-term memory to construct an overall, global structure of the text, which involves what Kintsch has called long-term working memory. Readers relate their preexisting knowledge, goals

and interests to the text as much as they interact with text-derived features. Thus, to learn from text, learners need to form coherent mental representations of the topic, both coherent in terms of the textbase and their prior knowledge.

Nassaji (2002) has extended Kintsch's model of reading comprehension to second-language (L2) learners, arguing the a top-down, schema theory of reading comprehension, whereby a reader's knowledge of the world provides the basis for comprehension, fails to explain why L2 readers with well-developed schema do not to read at the same rates and levels of understandings as first-language (L1) readers. Nassaji emphasizes the construction of the textbase, the representation of the text built in the process of comprehension, as a more important process for L2 readers than the application of their background knowledge. Thus, many L2 readers rely more heavily on linguistic ability if they do not share the appropriate background knowledge, which likely taxes readers' working memories and decreases their comprehension. However, research has also shown that L1 and L2 reading is interdependent (August & Shanahan, 2006; Fitzgerald, 1995), so L2 students may benefit from some transfer of reading proficiency skills in their first language.

What these cognitive processing models of reading have in common is an emphasis, if only implied, of the importance of language knowledge for successful reading comprehension. Word- and sentence-level features are integrated with the reader's pre-existing knowledge, linguistic and conceptual, to better understand the text. This language-based model of reading comprehension is elegantly summarized in Hoover and Gough's (Hoover & Gough, 1990) model of reading. Their Simple View of Reading summarizes early reading research to identify two key processes necessary for reading: decoding x linguistic comprehension. With recent research, we can posit that the linguistic comprehension of the learner is a compilation of language knowledge

such as vocabulary breadth (e.g. Nation & Snowling, 2004) and depth (Proctor, Silverman, Harring, & Montecillo, 2012), semantic knowledge (Ouellette, 2006), and syntactic knowledge (Huttenlocher et al., 2002). Hoover and Gough (1990) suggest that "instruction facilitating linguistic comprehension should … facilitate reading comprehension" (p. 153). Thus, how linguistic comprehension is developed becomes a key issue in this model of reading.

Currently, this relationship is increasingly understood to rest on a sociocultural or socio-constructivist theory of language learning, that interaction and engagement with others facilitates language development (Vygotsky, 1986), which, in turn supports literacy skill. Research into second language acquisition explains this theoretical position by emphasizing input (Krashen & Terrell, 1983; Long, 1981) and output (Swain, 2005) for linguistic competency to develop. Thus, to develop linguistic comprehension skills to support reading skill, the learner should have rich opportunities for linguistic input and output, which necessarily means interaction between the learner and others.

Social perspectives on the reading process focus on how language functions in dialogic ways between the reader, the text, and other actors in the context of a reader's activity. These perspectives on reading describe it as a "dialogic exchange of meaning or transformation of mutual knowledge between writer and reader mediated by the text" (Nystrand, 2006, p. 396), a context dependent process that relies on interactions between not only the reader and writer, but the teacher, classroom context, and the text itself (Ruddell & Unrau, 2004). This perspective emphasizes the environment around reading as a key aspect of the reading process, interacting with the knowledge construction process, where comprehension is affected by motivation and stance, as well as language and processing strategies (Ruddell & Unrau, 2004). Similarly, Rosenblatt (1994) emphasized the dialogic nature of reading: like in a conversation, the reader

comes to the text with a unique linguistic and experiential reservoir from which the reader draws meaning. Thus the same text may seem different to different people. Readers adopt stances toward text as they interact with texts through the processes of selective attention, the linguistic experiential reservoir, and the development of background knowledge as they read. Rosenblatt argued that this stance falls on a continuum of efferent to aesthetic, whereby the reader comprehends and treats the text as either features to be extracted and retained or as a lived experience, attuned to the more emotional and aesthetic features of reading.

This brief review of two main theoretical accounts of reading are pulled together in this review in service of employing the Simple View model of reading to frame how reading comprehension is conceptualized in this study and how it can explain the importance of language input and output in elementary classrooms. The social and cognitive aspects of reading comprehension suggested in these models of reading are elements of Hoover and Gough's (1990) concept of linguistic comprehension in their simple view of reading model. While these linguistic views on reading processes were not altogether formulated when Hoover and Gough published their model, research from both perspectives can shed light on how linguistic comprehension develops for a reader to read with comprehension. Linguistic comprehension is also mitigated in part by the context of reading. Before moving to a discussion of how researchers have explored pedagogical discourse as a support for reading, we briefly summarize the reading context for young readers as they develop language and reading skills.

The curricular context of reading comprehension.

The reading research community has established the necessity of oral language skills in early literacy development (Snow, Porche, Tabors, & Harris, 2007), and there is broad consensus that dialogic interactions can promote complex language skills to access texts as they get

increasingly more demanding and abstract (Bronfenbrenner & Morris, 1998). As students enter middle school, the range and complexity of variables affecting reading skills change. Successful reading not only entails word-reading skills but "also knowing vocabulary, being able to process longer and more complicated texts, having the background knowledge presupposed by the text and the curriculum, and having some reasonable level of interest in the text and motivation to persist in reading it" (Snow et al., 2007). This is the nature of the expectations that shift from when students move into reading to learn. Around fourth-grade in US schools, "texts and other materials typically read become ever more varied and complex in content, language, and cognitive demands. In order to read, understand, and learn from these more demanding texts, the readers' knowledge, language, and vocabulary need to expand, as does their ability to think critically and broadly" (Indrisano & Chall, 1999). This has important consequences for not only successful reading outcomes, but successful learning and school outcomes as well. The development of reading comprehension, as expectations and text change, is therefore one critical issue of student achievement.

Decoding and building fluency are foundational reading processes learned in the early stages of literacy acquisition (e.g. Connor, Morrison, & Slominski, 2006) but by fourth grade, students are expected to become independent and strategic readers and use their skills to engage with text to analyze, synthesize, organize and evaluate (National Council of Teachers of English, 2006b). As texts become more decontextualized and abstracted from the everyday lives of readers as they develop academically in schools, the success of reading relies more heavily on the language of the text and less on the context (Snow et al., 1991). Decontextualized language tends to use richer and varied vocabulary (Nagy & Scott, 2004), and more complicated syntactic structures. By the upper elementary grades, academic work presupposes literacy skills involved

in most learning activities (Snow et al., 1991), which places upper elementary and adolescent readers in a unique position compared with their younger peers: as readers, they not only have to rely more on their language knowledge, but they must know more language that is not readily available in everyday contexts.

Therefore, it is in this schooling context that frames the current study. In the upper elementary grades, the shift toward a reliance on academic, decontextualized language skill demands skilled and responsive teaching to meet the changing demands of comprehension. After all, language is a system of choices, and how language users exploit and employ language to make meaning is a crucial feature of pedagogical support (Christie, 2002).

Instructional Discourse and Reading Comprehension

The contexts of reading comprehension across grade levels clearly paint a dynamic picture of the demands placed on readers. Therefore, teaching for high-level comprehension need to be dynamic, flexible, and contextualized (Wilkinson & Son, 2011). In this section, intervention students and reviews of instructional programs that have attempted to address these concerns are reviewed.

Early reading instructional programs that have focused on the explicit teaching of comprehension strategies have shown positive effects on students' reading achievement. For instance, Palinscar and Brown's (1984) Reciprocal Teaching approach, emphasizing the comprehension strategies of summarizing, questioning, clarifying, and prediction, was found to have effect sizes ranging from a median of .32 to .88 on student reading outcomes (Rosenshine & Meister, 1994). Questioning the Author (Beck, McKeown, Sandora, Kucan, & Worthy, 1996) has focused more on the content of the text and meaning-based questions (Wilkinson & Son, 2011). Both instructional programs draw from a constructivist perspective of reading, instructing

students to actively make sense of the text. Questioning the Author and Instructional Conversations (Saunders & Goldenberg, 1999) also emphasize text-based discussions, which emphasize the more dialogic approach to reading in the same vein as the work of Wells, (1990), Nystrand, (2006), Rosenblatt (1994) and Gee (2001), discussed above. Wilkinson and Son (2011) suggest in their recent review that a "fourth wave" of comprehension instruction reflects a more dialogic perspective, emphasizing content-rich instruction, discussion, argumentation and intertextuality. Instructional programs such as Collaborative Reasoning (Waggoner et al., 1995) and the Paideia Seminar (Billings & Fitzgerald, 2002) are explicit in the centrality of text-based discussions for reading comprehension, and have evidenced the shifting roles of the teacher from interpretive authority toward a more facilitative role in discussion (Billings & Fitzgerald, 2002; Chinn et al., 2001), while work in Collaborative Reasoning has also linked this instruction to reading and writing outcomes. Small group peer-led CR discussions have been found to affect written argumentation (Reznitskaya et al., 2001), and another study found that students who participated in CR groups scores significantly better than the control group on the cloze reading tests (F= 7.36, p = .009), as well as on listening comprehension and writing outcomes (Zhang, Anderson, & Nguyen-Jahiel, 2013).

Many of the more pedagogically popular reading comprehension instructional programs were subject to two recent syntheses, a meta-analysis (Murphy et al. 2009) and a study validating this analysis (Soter et al., 2008). These two papers reviewed the effects of nine classroom discussion programs on the development of literacy skills (including Collaborative Reasoning, Questioning the Author, and Literature Circles) and provide insight into the instructional practices that correlate with this development. Murphy et al. (2009) found that many of the nine discussion programs were effective at promoting literal and inferential comprehension, as well as

effective at increasing quantity of student talk. However, few approaches promoted development of critical reasoning or argumentation, though these findings were likely attenuated by the research designs employed by the studies reviewed. Instructional programs that emphasized student control over discussions emphasized an aesthetic stance (Rosenblatt, 1994) to reading comprehension, while those programs in which teachers were more in control tended toward a more efferent stance. In contrast to some of the assumptions in these programs of instruction, Murphy and colleagues found that the quantity of teacher talk was not inversely related to comprehension gains, because less teacher talk did not result in improved student comprehension.

Soter et al. (2008) qualitatively reviewed the same nine discussion programs to better understand not simply whether talk can enhance text comprehension, but more specifically which qualities of teacher talk relate to students' comprehension and critical thinking skills. The authors assessed the nature of questions, the presence of elaborated explanations and reasoning words, and exploratory talk. They concluded that the most productive discussions for high levels of comprehension and thinking were instructional periods when students controlled the discussion frequently, where discussions reflected increased student uptake, and where teachers and students asked authentic questions. Both the Murphy et al. (2009) analysis and the Soter et al. (2008) synthesis found a wide range of grade level samples in the research, but the most commonly sampled for these intervention studies were with fourth- through sixth-grade students. The interventions reviewed in both works were also quantitative research studies, which revealed positive correlations between the discussion programs and student reading achievement outcomes, but were limited in describing these dialogic contexts, and how they might be supportive of students' text processing and comprehension skills.

There is evidentiary support for the claim that oral discussions also benefit language-minority students. Saunders and Goldenberg (1999) compared four treatment groups, consisting of variations in the components of instruction, like the inclusion of literature logs only or logs and instructional conversations, to a control group. Those students involved in the instructional conversation showed statistically significantly higher results on factual and interpretive comprehension tasks. English learners benefitted the most from instruction that used both literature logs and instructional conversations. Klingner & Vaughn (2000) examined reading activities during science instruction in one class of 37 students (35 students spoke Spanish as their L1) for peer-peer helping behaviors. The analyses showed that the majority of students' helping behaviours (43%-56%) addressed unknown words in the texts, in contrast to 15%-34% of their discussion time spent on "getting the gist" strategy. For bilingual students, clarifying a peer's understanding meant they first had to ascertain if their peer knew the concept so that they could simply translate the word into Spanish, as well as decide if the word was important enough to warrant an extended explanation, or if a simple translation in to Spanish would suffice. Students' posttest scores were statistically significantly different than their pretest scores, and student gains were higher on the second chapter of the reading discussion group. The authors suggest that while the test score gains were modest, these scores may indicate that receiving and giving help through collaborative reading strategies benefited these students.

The literature on classroom discussion and bilingual students includes the additional issue of language choice during discursive activities that research on monolingual classrooms does not. These works also indicate similar attention to the same units of analyses in 'monolingual' studies, including attention to the form and function of questions, quantitative counts of types and tokens, and the positioning of students in relation to academic discourse and content

instruction. However, more studies are needed to better understand the effects of exposure to, and the participation in, instructional discussions around texts in both English-only and bilingual samples. Since English learners continually play 'catch up' to their English-only peers in most areas of language and literacy development, discussions would seem to play a important scaffolding role in textual understanding and reading skill, as evidenced in the Saunders and Goldenberg (1999) and argued by other researchers (e.g. Gutiérrez & Rogoff, 2003).

Nystrand (2006) drew two main conclusions in his review of research on the relationship between classroom discourse and reading comprehension: 1) studies from cognitive, sociocognitive, sociocultural, and dialogic perspectives "strongly support the potential of classroom discussion to enhance reading comprehension instruction," and 2) recent research emphasizing the dialogical and sociocultural nature of discourse indicates that discussion works as an epistemic environment that supports literacy development (p. 401). Nystrand reviewed conceptual work spanning 1860 to present day, noting that educational researchers have been well aware of the widespread use of recitation in US school lessons, but it is only since the 1970s that research began measuring the effects of classroom discourse on students' reading comprehension (Christie, 2002). Among his conclusions, Nystrand notes the evidence supporting the positive impact of small-group discussions on individual reading comprehension versus whole-group discussion, but also suggests that there is more evidence of positive outcomes for student reading comprehension when involved in whole-class discussions, as long as these discussions included envisionment building (Langer, 1992), the integration of reading and writing with talk (Nystrand, Gamoran & Carbonaro, 2001), and features of programmatic talk, such as questioning, summarizing, and turn-taking. Thus, some important features of talk for reading comprehension are "the questions teachers ask, how they respond to their students, and

how they structure small-group and other pedagogical activities" (Nystrand, 2006, p. 400). However, his review offers few hypotheses to better understand why it is that this sort of classroom environment may support students in reading text with a depth of understanding needed for effective and successful academic work.

Two other recent reviews of research on oral discourse and literacy outcomes provide similar conclusions. Slavin and colleagues (2008) conducted a synthesis of approaches to improving reading in middle and high school populations, and found that approaches to professional development that privileged changes to daily teaching practices, such as implementing cooperative learning in small groups and with technology, had positive reading achievement effects. In one study, Pressley et al. (2001) showed the most effect literacy teachers for their contexts in first grade used their knowledge of students' abilities to scaffold, provided positive reinforcement, and made connections between reading and writing across the curriculum. Pedagogical knowledge for literacy development includes not only the consideration of activities and tools, but how language is used a tool to scaffold the appropriation of literacy skills.

The most recent review of research synthesizing work on classroom discourse and literacy development has continued to find support for this relationship. Lawrence and Snow (2011) reviewed correlational and quasi-experimental studies linking classroom discourse to literacy outcomes, and summarized the findings on teaching moves effective for conducting rich discussions and the qualities of rich discussions. The authors listed four teaching moves that research has shown to be a focus of research over the last three decades, including modeling, direct explanation, marking, and revoicing. Their review of intervention studies suggested that successful discussion programs shared features of 1) increased student interpretation of texts, 2)

a focus on 'air time', 3) who controls the topic under discussion, and 4) the inclusion of open-ended questions. They therefore concluded that the qualities of rich discussions include: open-ended questions, students and teachers sharing talk time and interpretive authority, opportunities for peer-to-peer interaction, the establishment of a goal for discussions, and explicit attention to discussion rules. This review also found a dearth of research on how, when, and why discussion contributes to the development of literacy skills and content learning. For instance, the authors noted that while it seems plausible that features of rich discussion in secondary classrooms would be workable in elementary classrooms, there are no systematic studies exploring students' developmental trajectories and capabilities in participating, and benefiting from, rich discussions. Furthermore, the literacy field's understanding of the beneficial features of discussion is still nascent. And literacy researchers continue to know little about effective implementation of discussion programs and protocols that lead educators to conduct rich discussions for their learners.

The Nystrand (2006) and Lawrence and Snow (2011) reviews indicate that research on classroom discourse that has been studied in two distinct ways: using discourse analysis to describe linguistic utterances by participants, and correlational studies that look more generally at discussion-based pedagogy for reading instruction and its predictive power on student outcomes. Studies on classroom discourse and discussion have predominantly used discourse analysis methodologies, necessitating small samples and detailed linguistic data. Applebee and colleagues (2003) have argued that statistical modeling that accounts for students clustered in classes and schools is a promising methodology to capture the effects of instructional discourse for reading comprehension in larger samples. The notion of an effective methodology to study

learning outcomes and instruction becomes important as the field asks more questions about approaches in teaching for in-depth understanding (Bransford et al., 2000).

Dialogic Features of Effective Literacy Instruction

The present research study takes a microgenetic look at classroom discourse for specific features of instruction that relate to reading outcomes. As has been discussed above, discussion-based reading programs have shown generally positive effects on student learning and literacy outcomes. Open-ended questions, interpretive authority for students, and peer-peer interactions are some of the features shown to have positive effects (Murphy et al., 2009). In the following sections, evidence for specific instructional discourse moves is reviewed as they relate to literacy outcomes. These features form the basis of the coding scheme used for the current study, described in detail in Chapter 3.

The study of discourse and literacy outcomes in secondary classrooms has tended to explore unprogrammatic instruction involving academic discussions, while research in elementary classrooms has looked at the results of discussion-based reading programs. We first present findings from secondary classroom contexts, and then turn to primary contexts in this section.

Findings from secondary classroom settings.

From these studies in secondary settings, features of effective language arts instruction that produce favourable learning outcomes include authentic questions that involve reciprocity of interaction between students and teachers (Nystrand & Gamoran, 1991a), or "dialogic instruction," comprised of open discussion, authentic teacher questions, and student uptake from questions (Applebee et al. 2003).

Using multilevel analysis, Applebee, Langer, Nystrand and Gamoran's (2003) built on the four authors' previous work in secondary classroom contexts, integrating a variety of variables into their analysis of the literacy performance of 974 students in grades 7, 8, 10 and 11. The researchers measured 64 classrooms across 19 schools, focusing on discussion-based approaches to investigate whether they were more successful at improving students' spring literacy performance, controlling for fall performance. Lessons were observed for the questions posed by teachers or students and whether there was open discussion (a free exchange of information between three participants lasting 30 seconds or more) in instructional episodes. Instructional materials and connections between activities were also noted. Overall results indicated that spring performance on the researcher-developed written outcome measure was associated with high academic demands and discussion-based approaches ($\beta = .15$, $p < .001$; $\beta = .21$, $p < .001$, respectively), controlling for background variables like SES, race, and the fall literacy measure. This was a main effect across both levels of schooling (middle and high), school location (rural, suburban and urban) and track. In all, this research study was powerful in its large-scale and widely-dispersed sample. Though it did not provide data on some actual discussion episodes to better understand how questions, uptake, or envisionment might sound during instruction, and the outcome measure was not reading comprehension, the findings are important for analyzing the effects of discourse moves that established episodes of discussion and related activities "designed to involve students in the exploration of ideas", a likely "powerful tool for learning" (Applebee et al., 2003, p. 722). This study also underscores the lack of such large scale studies looking at dialogic instruction and reading comprehension.

The Applebee (2003) study built on Nystrand and Gamoran's earlier work which indicated similar findings in similar sample populations. One earlier study looked at instructional

discourse as an indicator of substantive student engagement, measured as the authenticity of teacher and student questions, uptake, and the level of evaluations of students' contributions. These features were regressed on a curriculum-based reading and writing test in 58 classes (Nystrand & Gamoran, 1991). Substantive student engagement was positively correlated with literacy achievement, leading the authors to note the paradox that students achieve more when teachers attend to what students do know. This pedagogical attention was reflected in higher numbers of authentic questions rather than known-answer questions. In contrast, small-group work, which is suggestive of opportunities for students to share their knowledge, was negatively associated with achievement. The authors suggest that this was likely due to the nature of the tasks during small-group work, which were not collaborative but instead worksheet tasks.

Similarly, in a smaller study (N=16 schools), Gamoran and Nystrand (1991) used a series of regression models to explore the effects of background and instructional variables on student achievement in eighth-grade ELA and Social Studies classrooms. They found that lecture, question and answer, and discussion patterns all significantly predicted achievement scores in both content-areas. Race interacted with instruction in that Hispanic students' lower scores were made non-significant by activities and discussion in ELA classes. The study also indicated how rare discussion was in the sample, averaging less than a minute per day, confirming earlier research (e.g. Cazden, 1988). This finding is particularly interesting in relation to the authors' other work because in this study, discussion was defined rather vaguely as "when classroom discourse transcends the typical question/response/evaluation sequence to exhibit free-flowing comments among students and the teacher" (p. 281). Discussion was not defined in terms of any instructional discourse markers, only in contrast to question-answer recitation sequence (IRE).

Discussion was also considered in this article as an instructional activity, not a kind of instructional discourse.

Nystrand and colleagues refined and clarified their analysis, and in their 1997 study, questions became a key instructional variable in determining "general dimensions of instruction…central to instruction and regulat[ive] of the extent to which teacher-student interaction can be dialogic" (Nystrand et al., 1997, p. 36). Questions were coded for their source (teacher-student), their response (yes-no), authenticity (whether they presumed prespecified answers or not), uptake (previous answer incorporated into a subsequent question), and the notoriously difficult (Cazden, 1988) categorization of cognitive level (from questions eliciting recitation to speculation and analysis). Nystrand and his colleagues confirmed their earlier findings, that eighth- and ninth-grade classrooms (N = 112) were overwhelmingly monologic: teachers talked for all but an average of 50 seconds per class. Discussion in the eighth-grade sample showed a large effect on student achievement, a surprising finding since there was so few moments of discussion.

This body of work by Nystrand and colleagues constitutes the most systematic large-scale investigations into secondary discourse practices in literature classes. It is persuasive in terms of key instructional variables that point to dialogic discourse patterns, and consistently shows positive effects of discussion (however vaguely defined in early work) on student literacy achievement. The correlational nature of this work allows analysts to focus on the real-world effects of student literacy achievement, and still allows some attention to more discrete linguistic characteristics of instructional talk. However, as with all correlational work, it cannot attribute causes for why these sorts of questions, uptake, or discussion seem to support reading and writing at this grade level.

From this review of dialogic instruction and its effects on secondary students' literacy achievement, it can be summarized that the features of questions, uptake, recasting, and aligning students' current understandings with curricular content correlate positively with literacy achievement in secondary settings.

Findings from primary classroom settings.

With such consistently positive findings in secondary settings, one wonders if this positive relationship between instructional discussions and literacy and content outcomes is replicated in younger readers. Most studies of discussion-approaches and dialogic instruction at the elementary level have been conducted in math and science contexts (Chin & Osborne, 2010; Franke et al., 2009; O'Connor, 2001; Riley, 1986; Turner, Meyer, Midgley, & Patrick, 2003; Wiebe Berry & Kim, 2008), while few have included descriptions of naturalistic instructional discourse in language arts classes. Unlike the secondary studies, however, at the elementary level more work has been conducted on specific interventions designed to promote teacher and student discussions around texts (i.e., programmatic instruction) rather than naturalistic observations and correlational analysis. There is a body of work that has looked at instructional impacts on beginning readers that can be considered language arts contexts, but the early contexts of learning to read differ from the emphasis on reading for understanding in upper elementary literacy contexts. As such, dialogic instruction has not been the focus of these early studies; instead they have focused mainly on instruction of word-level reading skills like the alphabetic principle, decoding, word fluency and vocabulary knowledge (Aukrust, 2007; Connor, Morrison, & Underwood, 2007; Silverman & Crandell, 2010). However, these studies, along with the extant studies of discussion in upper elementary grades, are a window into instructional responses to the developmental of literacy skill over the elementary years. Research has

demonstrated that time spent in text-based discussions described in studies of secondary contexts can be conducted in fourth and fifth grades (Chinn et al., 2001; Eeds & Wells, 1989), and is likely productive in the early grades for building vocabulary (Aukrust, 2007; Silverman & Crandell, 2010) that supports reading comprehension.

Studies in the early years on the instruction for vocabulary to support reading development derive from the well-established connection between vocabulary knowledge and reading comprehension (Anderson & Freebody, 1981; Hart & Risley, 1995; Stahl & Fairbanks, 1986). Vocabulary plays an increasingly important role in later reading comprehension (McGuinness, 2005; Snow, Griffin, & Burns, 2005), and therefore is an important focus for reading research in the elementary grades. For example, a teacher's use of sophisticated and varied vocabulary and analytic talk in preschool settings has been related to students' fourth-grade reading comprehension (Dickinson & Porche, 2011). Silverman and Crandell (2010) found that explicit instruction on words during read-aloud time showed a positive main effect on vocabulary performance, as did contextualization of words during non-read-aloud time. In a study of four first-grade classrooms, differential effects of instruction were found on students' reading skill, including students with early low literacy benefitting more from early phonics and later vocabulary instruction and text-based discussions (Juel & Minden-Cupp, 2000).

In sum, studies in naturalistic classroom discourse at the upper elementary levels are sparse, as are studies correlating discussions used during instruction with student literacy outcomes. While some discourse analytic studies at this level suggest some important features of dialogic talk, such as student management of discussions, exposure to explanations, and the quality of questioning, it is unclear how the demands of reading to learn plays out in the instructional talk in these classrooms, and if dialogic instruction has an effect on reading

comprehension. The research on Collaborative Reasoning suggests the possibility of positive correlations between participation in dialogic instruction and reading. However, reading comprehension has yet to be a well-established outcome variable in such studies of dialogic instruction, nor linked to discreet features of dialogic instruction lauded in the previously reviewed studies. In the following sections, evidence for specific talk moves is reviewed briefly in order to describe and contextualize the coding scheme used in the present study. These features of dialogic instruction are briefly reviewed below, and provide the framework for the current study's analytic plan.

Teacher questions.

In her seminal work, *Classroom Discourse*, Cazden (1988) notes a common interest in the research community in analyzing teacher questions since they are prevalent in typical instructional contexts. Consequently, the research literature has established the importance of questions in classroom talk, studying the impact of questions as "sites of interaction" on dialogic episodes (Nystrand et al. 2003) and the role questions play in students' construction of knowledge (Chin & Osborne, 2010; Graesser & Olde, 2003; Scardamalia & Bereiter, 1991). Questions have been categorized most commonly as open or closed, authentic or inauthentic, or high- or low-level cognitive questions. Both teacher and student questions, if authentic, have been shown to be significantly related to discussion, "the unprescripted exchange of ideas among students and the teacher," with student questions showing a large effect on dialogic episodes during secondary literature instruction (Nystrand et al., 2003, p. 185). Research in high school science classes has shown similar findings to the work in language arts contexts, indicating the importance of teacher questioning and explicit uptake (discussed below) during scientific discussion. Though dialogic interactions were studied in only three classrooms, McNeill and

Pimentel (2010) found that open questions likely pushed students to more robust argumentation and increased the amount of student talk during science lessons, an important finding as science curricula emphasize highly-skilled argumentation.

The epistemological assumptions grounding questions does seem to matter beyond whether a question is open or closed in its intention. Authentic questions have been a consistent focus of Nystrand's work, who drew on early classroom discourse studies that addressed the nature of inauthentic questions during IRE patterns, and those (rare) authentic questions that functioned to allow students more freedom to express their contributions (Nystrand, 2006). Authenticity in querying is suggestive of a different epistemological stance than a pattern of talk characterized by posing inauthentic question (Nystrand, 2006). In an instructional context, a teacher asking students questions without presupposed and prespecified answers emphasizes the importance of student's thinking and verbalizing their process of understanding. This is a more constructivist view of knowledge making, as opposed to establishing a pattern of display questions (Boyd & Rubin, 2006) that assess student knowledge more so that engage with it. In this sense, authentic questions are necessarily open questions, since authenticity implies that the questioner does not have a prescribed or known answer in mind, and this allows for a broader range of responses, lending an 'openness' to the talk. Questions provide a "powerful index of the extent to which teachers open their classes to student voices when they ask the questions they do" (Nystrand et al., 1997, p. 37).

This dichotomy has been challenged, however, by evidence that display and authentic questions attend to different facets of a text-based talk. In Boyd and Rubin's (2006) study of a small 4/5 class of English language learners, they noted that authentic questions attended mainly to students' opinions and background experiences, while display questions referenced the text

and generated more student talk that was extended, coherent, and socially engaged. For this group of language learners, it seemed that questions grounded in a text gave students more opportunities for output in their second language, an important tenet of second language acquisition theory (Swain, 2005).

The research on teacher questioning has been extensive, and figured prominently in process-product paradigm in the 1960s and 1970s, and more recently in sociolinguistic work in classrooms, emphasizing the social context of language over links between teacher behaviors and student outcomes (Carlsen, 1991). Important for this research is the recognition that teacher and student questions are reciprocal and context dependent (Carlsen, 1991), and therefore simply coding for questions misses the intent or the implications of questions for dialogic instruction. Furthermore, judging the type of teacher question is dependent on the context of the lesson in which the questions occur, in that the genre of the instructional discourse established by the instructor mitigates the nature of his/her questions (Nystrand et al., 2003). For instance, the question, "What were the causes of the Civil war?" in a review lesson is likely recitational, but the same question posed in a seminar would be more authentic and open-ended. Therefore, judging the type of question depends on understanding the genre of the lesson, as well as how the question unfolds in the proceeding utterances.

Uptake Questions.

Taking up student contributions encourage a reciprocity in the patterns of instructional talk. The notion of uptake has been used as a covariate predicting dialogic instruction (Applebee et al., 2003; Nystrand & Gamoran, 1991; Nystrand et al., 2003). Uptake questions occur when one participant asks another about a previously stated idea (Nystrand et al., 2003). It is an indicator of the reciprocity of talk on a topic, or the "contingency" of questions on preceding

utterances (Boyd & Rubin, 2006). Others have characterized this feature of instructional discourse as connected interactions that support, refute or question a previous idea (McNeill & Pimentel, 2010). Uptake is commonly marked by pronouns referring to the preceding comment: these deictic references (Nystrand et al, 2003, p. 146), such as 'Why do you think *that*? Or 'how did *it* work?' make the student's contribution the "momentary topic of discourse" (Nystrand et al. 2003, p. 146), thus playing an important role in the reciprocity of understandings and meanings since the student and teacher listen and respond to each other. This may be a second key aspect to uptake questions: posing them may encourage student attention to the learning activity, which, in the early years, has been linked to reading outcomes (Dickinson & Porche, 2011). Uptake questions also lend a sense of contiguity to the unfolding lesson discourse, providing some coherence to the discourse as well as providing opportunities for students' articulation of their understandings. McNeill and Pimentel's (2010) data revealed one teacher's use of uptake may have helped students to use consider the views of their peers and reflect on their thinking more readily, supporting scientific argumentation, an increasingly emphasized goal of science education. In other work, uptake as a covariate showed a significant effect on discussion episodes, defined as an unprescripted sharing of ideas (Nystrand et al., 2003).

Cognitive level of questions.

Categorizing and rating the cognitive difficulty of pedagogical questions may aid educators, as seen by the influence of Bloom's Taxonomy in teacher preparation courses and materials, but may also be imprecise for research purposes (Cazden, 1988). Despite this warning, researchers have operationalized the cognitive level of instructional questions with some success (e.g. Nystrand et al., 2003). A sense of the quality of teacher questions posed during instruction around texts can indicate the teacher's stance in relation to the text, and to the act of reading

more generally. Is a text treated as a resource to obtain facts, or a resource to engage with critically, for instance? Reading in the upper elementary grades demands more generalization, analysis, speculation, and evaluation of texts, in order for more in-depth understanding of content, and teacher questions can structure these opportunities for students to develop these reading skills. Higher-level questions that elicit such thinking "open up the cognitive field beyond a mere reporting or replication of another's voice, incorporating the possibility of the speaker's added perspective and particularity," (Nystrand et al., 2003, p. 147) forwarding the lesson through student engagement rather than a recitative pattern of known-answer questions. Nystrand and colleagues (2003) coded for lower and higher level questions like generalizations, analysis and speculations and found that high cognitive level question were significantly correlated with transitions of instructional talk into discussions. Interestingly, this study found that high cognitive-level questions from teachers tended to stimulate discussion, while at the same time discouraging student questions.

Teacher evaluations of student responses.

Cazden (1988) and others (see Nystrand, 2006; Nystrand et al., 2003) have argued that analysts need to look at longer sequences of discourse, not only the question that typically begins instructional discourse, since it is frequently the case that the observer cannot understand the nature of a teacher's question without a student's response and the teacher's evaluation (Cazden, 1988; Nystrand et al., 2003). Thus, longer sequences of classroom discourse can be seen as scaffolds for learning, with questions typically constructing the initial scaffold, and "reconceptualizations" of a student's given response as a follow-up scaffold, ideally giving the student an opportunity to conceptualize the topic in a new way, or connect to previous understandings (Cazden, 1988, p. 111). Teachers' evaluations of student contributions can work

students' answers into the lesson discourse, or modify the topic to provide depth to the topic at hand to provide scaffolds for student understanding. Teachers' responses to students' contributions can also preclude pushing student thinking or establish a pattern of discourse that is recitational rather than dialogic.

Similar to Cazden's work (1988), and Nystrand and colleagues (1997) write that though discourse encompasses much more than questions and answers, the study of questions and answers allows analysts to look at general dimensions of instruction since questions and answer episodes are so central to instruction and regulate the extent to which teacher-student interaction can be dialogic. Questions presume answers, so they are "negotiations of sorts," revealing features of teacher- student interaction. Nystrand and colleagues (1997) have argued that the degree of authenticity, uptake and the level of cognitive activity that questions elicit are key to understanding this reciprocity (or lack thereof).

Evaluations of student contributions can also be evidenced by teacher "revoicing," a powerful discussion tool that O'Connor and Michaels (1993) identified in two elementary math classrooms that were characterized by frequent discussions of math problems, though this dialogic feature is not limited to math instruction by any means. Through detailed discourse analysis of these discussions, the authors noted a consistent use of connections made by the teachers to previous speakers and to the content of the lesson. For example, the authors argue that the teacher utterance, "So then, you don't agree with Sarita that if she picks a number halfway between…" acts in three ways: 1) to allow the teacher to make explicit her inference as it relates to the content under discussion, 2) to allow the previous speaker (a student) a chance to validate the teacher's inference, and 3) to position that student as involved in the participant structure of the classroom (O'Connor & Michaels, 1993, p. 322). This responsive talk move

allows the teacher to clarify or alter a student's contribution to scaffold the discussion for students, attending to students' sense making and understanding of the content at the same time as addressing principles of democratic education by involving students' in the reasoning of the discussion (Michaels et al., 2008).

A final note about teacher evaluations of students' contributions is warranted. Particularly important for creating responsive pedagogical environments for multi-cultural and multi-lingual classroom contexts, scaffolds, reconceptualizations, and revoicing should be drawn from students, and additive so as not to deny the meanings students bring to school (Cazden, 1988). In the field of second language acquisition, for instance, research has shown that recasting of student output benefits higher-proficient English learners' morphology and syntactic knowledge (Mackey & Philip, 1998), and when language learners understand recasting to address linguistic form and not meaning (Nicholas, Lightbown, & Spada, 2001).

Nystrand and colleagues (2003) judged teacher evaluations to be high only when the teacher incorporated a student's answer into the evaluation (p. 147). In our data, it was much more common for teachers to validate the student's contribution and further it with some scaffolding, like providing a synonym or spelling during a word study lesson, or extending the student's answer to link it to a previous lesson, reminding students explicitly of previously learned material. Nystrand and colleagues (2003) argue that high level evaluations set up the student as a source of new information, but in our data, it was more so the teacher that established and carried out this role. This could be an important difference between the grade levels of my sample and Nystrand and colleagues work in high schools.

Explanations.

Explanations have been shown to support younger children's vocabulary growth (Aukrust, 2007), but at the higher levels, it is unclear whether teacher explanations around texts act as direct instruction that can discourage dialogic patterns of talk, or if they serve as scaffolds that can bring students into conversations by supporting their developing understanding of a topic. Soter et al. (2008) have suggested that *student* explanations are indicators of elaborations and thus increased student talk, which has shown to be associated with reasoning and high-level thinking.

Studies that have investigated teacher explanations and direct teaching have shown some positive outcomes for student reading comprehension when direct instruction of reading strategies was a focus of instruction (Baumann, 1983, 1984; Pressley et al., 1992). An early definition of direct instruction was instruction having "an academic focus, precise sequencing of content, high pupil engagement, careful teacher monitoring and specific corrective feedback to students" (Duffy & Roehler, 1982, p. 35), which clearly does not preclude students from instructional talk, but also does not suggest any reciprocity to it either. The early literature on direct instruction and reading comprehension focused on explicit instruction of reading strategies to ameliorate the traditional tendency of teachers to ask assessment questions rather than "create an understanding of either the main ideas or strategy for figuring out the main idea" of a text (Gersten & Carnine, 1986, p. 71). These ideas of direct teaching were extended to include more jointly determined "transactions" around texts (Pressley et al., 1992), and later called into question for its focus on strategies over the actual meaning of the text (Pressley, 2000). However, this research has provided evidence that direct instruction of reading was an impactful support on students' identification of main ideas (Baumann, 1984; Stevens, Slavin, & Farnish, 1991), and

reading outcomes when part of a culturally responsive instructional program (Tharp, 1982). On the other hand, current research on kindergarten reading instruction using multilevel modeling has found that teacher demonstrations of reading skills were not related to early reading skills (Smolkowski & Gunn, 2012), though this finding may have been mitigated by low reliability scores and the limitation of measuring these talk moves as rate and not quality (Smolkowski & Gunn, 2012).

Aukrust (2007) and Rydland et al. (2013) explored the impact of exposure to teacher talk, particularly teacher explanations, on L2-speaking children's vocabulary acquisition. In the 2013 study, the authors found that teacher-led talk predicted the vocabulary differences at age five, and these differences held into grade five. In the 2007 study, twenty seven Turkish-speaking children in Norwegian preschool and grade school settings were followed for three years to map their vocabulary growth in both languages. Findings suggest that the quality of talk was important for L2 vocabulary acquisition, and that teacher talk appeared to have a long-term impact on children's vocabulary acquisition. Discourse analysis showed that teachers offered more word types and tokens within complex discourse, and demonstrated the interrelatedness of these three indicators when, combined into a composite variable, revealed no additive predictive power to the regression analysis.

L1 use: Unique dialogic opportunities in multilingual classrooms.

Second language acquisition theory has emphasized the importance of input and output in the second language acquisition process (Cummins, 1991; Krashen & Terrell, 1983; Swain, 2005; Swain & Lapkin, 1995), and the interdependence of bilingual readers' two languages (Cummins, 1979). Principles underlying the idea of dialogic instruction, and some models of reading, suggest that for successful literacy achievement, optimal learning takes place when

students can use their background knowledge and participate with others to construct their understandings through discussion, since the process of learning, understanding, and knowledge-building has been argued to be composed of social activities that occur in culturally- and instructionally-mediated contexts (Bransford et al., 2000). Multicultural and multilingual classroom contexts suggest a diversity of background, content, and linguistic knowledge that can aid to expand discussions by increasing the amount of divergent perspectives and thus opportunities for higher-level thinking skills like critical thinking. Multicultural and multilingual classroom contexts also suggest a complexity to this optimal learning context of emphasizing student knowledge and dialogic instruction. Teachers may have to scaffold more for dialogue to take place, for instance. One resource unique to bilingual students is the use of their first language, which in the current sample of this proposed study, is not the language of instruction. However, there are studies that suggest the importance of using students' first language to scaffold their understanding (Goldenberg, 2011; Klingner & Vaughn, 2000; Orellana & Reynolds, 2008), and thus their contributions to classroom activities such as discussions around texts. The use of a student's L1 to clarify a topic or student contribution to dialogic interaction during class may be a feature of dialogic instruction unique to multilingual contexts.

A multi-level analytic study of teaching effectiveness with bilingual kindergarten students looked at the instruction of 141 teachers teaching students to read in English and/or Spanish. Among the many findings of effective teaching for reading, the use of Spanish showed better outcomes than English use in the bilingual sample of children (Cirino, Pollard-Durodola, Foorman, Carlson, & Francis, 2007). In contrast, in a study that examined the impact of an instructional intervention on Hispanic and non-Hispanic K3 students, Gunn and colleagues (2005) found that Hispanic students benefited as much or more than non-Hispanic students from

supplemental reading instruction that focused on directly teaching word recognition, reading fluency, vocabulary, and comprehension skills in English. Studies such as these suggest that there are differences to be expected in reading outcomes when interactions between language status and instruction are examined.

Conclusions from the literature

By analyzing and quantifying discreet utterance features like question types and the quality of evaluations and questions to better understand dialogic practices in language arts lessons in upper elementary contexts, one can understandably be wary of losing a focus on the overall nature and patterns of literacy lesson discourse. However, it has been shown that dialogic episodes tend to occur in proximity to such features as uptake and questions eliciting high cognitive demands (Nystrand et al., 2003). This is not to say that more distal features of lessons, like the discursive genres or the textual materials the teacher chooses, do not also influence whether dialogic instruction happens, but that there is likely a relationship between these more distal, contextual features in the overall classroom discourse and the more detailed, utterance-level features that actually occur between teachers and students. In order to begin to capture the influence of discreet talk moves and more distal variables like student reading comprehension, a mixed analysis is used in the present study. Few studies in literacy research have used mixed method analyses to attempt to capture the micro and macro nature of language and literacy development in classrooms. In the following section of this literature review, extant mixed method research in discourse and literacy is reviewed to explore the power of attending to different levels of influence and context on student literacy achievement.

A review of dialogic instruction and its effects on literacy achievement in K-12 schools reveals that the great majority of the existing literature has focused on secondary and elementary

math and science settings. Neglected are language arts classes in the upper elementary grades where literacy curricula first begin to demand more complex language and literacy skills in advance of the more domain-specific teaching and learning that take place in secondary contexts. Also sparse are studies that used reading comprehension as an outcome variable. However, the work that has been done in this area has documented some promising features of dialogic instruction that may well support reading comprehension at the upper elementary grade levels. Thus, there is a striking need for research to explore the nature of teacher talk during upper elementary literacy instruction and examine its impact on the growth of student reading and vocabulary achievement. This is the express intent of the current research.

Review of the Methodological Literature

This final section of this chapter is a brief review of the methodological literature focused on classroom discourse, with a focus on the potential of mixed method designs to address some of the current limitations in this area of teaching and learning.

The global and specific aspects to studying classroom discourse.

The study of classroom discourse continues to be an important vein in educational research, not only because students in the developed world spend so much of their time in classrooms, but also because of the current scrutiny on teaching quality in high-stakes accountability contexts such as in the US. And as language is the primary mode of accomplishing schooling, we must engage with understanding these patterns of talk (Christie, 2002). Furthermore, since language is both the goal of literacy instruction and the medium through which literacy is learned, as well as the means through which research occurs (Bloome, Carter, Christian, Otto, & Shuart-Faris, 2005), attention to classroom language is crucial to understanding how learning, particularly literacy learning, may occur. Thus, discourse analysis,

in its many genres, has been the main methodological tool available to researchers over the last 30 years.

Classroom talk is a broad term that has been researched in two general ways. One line of research on classroom discourse links discourse to academic outcomes. This body of work ranges from quasi-experimental and meta-analyses (e.g. Murphy et al., 2009) to surveys and case studies, and has looked at discussion-based pedagogy and its effect on student outcomes (e.g. Applebee et al., 2003; Nystrand, 2006; e.g. Palinscar & Brown, 1984). A meta-analysis of the effects of classroom discussion on students' text comprehension and learning found that while many approaches are effective at promoting literal and inferential comprehension, few developed critical reasoning and argumentation, in part because the effects seemed to be attenuated by the study designs (Murphy et al., 2009). The effects of the various interventions were weaker in experimental and quasi-experimental designs than with single-group, pretest-posttest designs (Murphy et al., 2009). These studies tend to study macro variables like overall discourse patterns and achievement scores (Nystrand, 2006), and generally find positive effects on literacy when students are engaged in class discussions, and that these effects are academically measurable gains. However, these studies lack a description of the content of this language, an important component of understanding how instruction, delivered almost exclusively through an academic language register, scaffolds, frames, and influences students' learning.

To determine which features of classroom discourse may account for student achievement outcomes, it is therefore important to describe these features. A second approach to exploring classroom talk focuses on the content of language during classroom instruction, or the micro variables of language (Nystrand, 2006). These studies predominantly use discourse analysis to analyze classroom talk data. Discourse analysis has been defined as the analysis of

language to understand why and how language is used (Fairclough, 2003), or the analysis of "language-in-use" (Gee, 2005, p. 5). Sawyer (2006) uses the term "interaction analysis" instead of discourse analysis to refer to "all methodologies used to study verbal and nonverbal interaction" (p. 190). Each theorist emphasizes the language used in interaction with others, but the variety of units of analyses to describe the content, the form, or the function of instructional discourse has been large. Units of analyses in discourse analytic classroom talk studies have ranged from teacher and student questions (Applebee et al., 2003; Chin & Osborne, 2010; Nystrand et al., 2003), scaffolding techniques (Maloch, 2002) including discussion management (O'Connor, 2001), explanatory talk (Aukrust, 2007; Hufferd-Ackles, Fuson, & Sherin, 2004), and translation strategies (Orellana & Reynolds, 2008), among others. It has been argued that this body of research allows for a more nuanced picture of the dynamics of talk and activities (Nystrand, 2006). This nuance can provide not only a picture of possible factors involved in student outcomes, but also can help educators become aware of various discourse strategies mediate student learning (Sharpe, 2008). "[Teachers] can plan lessons which support a dialogic classroom environment in which student interact cooperatively with the teacher and each other as they construct new understanding that transform their conceptual understanding" (Sharpe, 2008, p. 146).

These two general approaches to studying instructional talk are underpinned by different sorts of questions and different understandings of the meaning of discussion-based approaches to literacy. Many discourse analytic studies draw on sociocultural frames to focus on participation structures and the reciprocity, scaffolding, or vocabulary used by classroom participants (e.g. Brice, Shaunessy, Hughes, McHatton, & Ratliff, 2008; Chin & Osborne, 2010; Hufferd-Ackles et al., 2004; Maloch, 2002; Nathan & Knuth, 2003; O'Connor, 2001). The assumption in such

studies seems to be that participating in class talk allows students to develop appropriate academic language needed for high-skilled literacy achievement, and encourages students to develop interpersonal skills appropriate to accomplish academic tasks. Studies that focus on correlating features of instructional talk with literacy achievement tend to draw upon both socio-cultural and socio-cognitive theories of reading, correlating discourse in the context of a social culture— the classroom— with reading skills.

For the proposed research study, these two research veins, correlational and discourse analytic work, are treated as complementary, with the potential to provide a robust and nuanced picture of classroom discourse in both its effects on student achievement, and the possibilities of why this might be so. Indeed, Nystrand et al. (2003) argue for a methodology that allows for the simultaneous study of macro and micro language features. This attention to both the nuances of discussion in classrooms, and the achievement results of students' participation in such pedagogical activities would seem to ameliorate the limitations of a correlational approach to the study of the effects of discussion. Nystrand (2006) notes that the correlational approach addresses "static macro variables" such as overall discourse patterns and pre- and post-achievement scores (p. 401). The power of qualitative discourse analysis is its ability to capture the "dynamics of unfolding activities," since language is not static, but lacks generalizability beyond the case study (ibid.). Nystrand (2006) writes that:

> Any given classroom interaction or instructional event (such as a question and its response) is shaped not only by the inertia, or chain, of immediately preceding interactions, but also by the interactions of previous lessons, as well as by institutional factors (such as tracking and social, cultural, and demographic variables including race, ethnicity, and SES). That is, every classroom interaction and event has a history with both global and proximal antecedents. In addition, the same given

> classroom interaction or instructional event shapes teacher and student expectations for subsequent classroom interactions and for factors transcending the class, such as student achievement (Nystrand, 2006, p. 402)

By analyzing the micro discursive processes of a classroom over time from such a sociocultural framework, the analyst can begin to understand the typical forms of dialogic instruction and what they reveal about teachers emphases during literacy instruction. Furthermore, if one understands cognitive development to be a process of meaning making with others (Halliday, 1993; Vygotsky 1978), then this meaning making with others must be revealed through discursive actions. Yet, these micro processes of discursive actions take place in classrooms subject to more macro forces like standardized assessments, curricula, and cultural expectations, to which researchers must also be focused. One early example is Erickson and Mohatt's (1982) classic microethnographic study of two classrooms, one in which a teacher struggled to adjust to students' cultural norms. The authors argued that the social structures in the community were a reflection of cultural assumptions of the community and directly related to the teaching practices of the teachers. This is but one example of work that focus on how broad structures of cultures get enacted in classrooms (see e.g. McQuillan, 1998). Many researchers have thus argued the importance of studying the global and specific aspects of instruction. Juel and Minden-Cupp (2000) suggest an "intermediate-level view" of the classroom, including "microlevel interactions" (teacher-student work on writing for sounds for example) and an analytic frame wide enough that it captures student achievement and broad activities like reading group organization and forms of literacy activities (p. 482). Lerman (2001) has characterized this methodological work on classroom discourse as "zooming" in and out, the research selecting a place for her lens, taking into account macro and micro situations in the complexity of teaching

and learning (Lerman, 2001, p. 90). The current proposed research is in the same methodological vein, attending to both global (eg. reading outcomes, reading materials and activities) and specific (eg. types of teacher questions and evaluations) aspects of dialogic instruction.

Mixed method studies of instruction and literacy.

Turner and Meyer (2000) note in their review of research on the instructional contexts of classrooms that studies must involve multiple variables to capture the complexities of these environments. The authors recommend mixed approaches to such studies in an attempt to answer "the how and why questions," which necessarily draws on both research methodologies (p. 80). Unfortunately for literacy research, there are few extant mixed method studies that explore classroom discourse and literacy achievement or development. Those literacy or instructional focused studies we did find emphasized the complementary nature of their quantitative and qualitative findings.

In an experimental study of undergraduate students' reading, students in the experimental groups were given relevance instructions before reading, and this resulted in slower reading and better retention (McCrudden, Magliano, & Schraw, 2010). Interviews post-experiment showed that instruction influenced readers' goals and choice of strategies to affect their processing and comprehension. In a seventh-grade classroom, a mixed analysis explored students' conceptual understanding of a science lab using a writing heuristic, a textbook write-up, or a traditional lab write-up (Hand, Wallace, & Yang, 2004). Quantitative and qualitative results indicated better conceptual knowledge with the first two scaffolds. The authors argue that this is evidence of the importance of writing to learn in this content area. Similarly, Goldenberg and colleagues (2005) found the use of quantitative and qualitative datasets to be especially complimentary across multiple studies of Latino readers. When their quantitative findings ran counter to their

hypotheses around literacy, that student book use at home was not related to literacy achievement, observational studies of parents and children reading at home showed the importance of the parents' cultural schema of a "bottom-up" emphasis on learning to read. These qualitative data explained the strong and positive relationship of kindergarten literacy achievement and phonics worksheets. In their multiyear work, the researchers were able to "test hypotheses about generalizable findings" and conduct nuance research into the routines of family life that influence children's literacy development (Goldenberg et al., 2005, p. 31).

Conclusions from the Literature

To summarize the work on classroom discourse and literacy development in US classrooms, the literature in secondary and elementary classrooms has used two general approaches to study the evidence for the importance of instructional discussions on reading comprehension. The correlational approach used with large samples, such as the work by Nystrand, Gamoran and colleagues, has drawn on sociocognitive frameworks to identify and test covariates in regression models. The theoretical argument supporting this work is that collaborative and reciprocal discourse features that occur in the social context of the classroom can be internalized and applied to individual reading and writing work demanded by teachers and the curriculum. This body of work has rarely looked specifically at reading comprehension, nor at reading achievement in the later elementary years when curricula demand a reorientation of reading skills, from an early elementary focus on building word-level reading skills and oral language proficiency, to text-level skills and conceptual understanding through reading (Lesaux & Geva, 2006).

The second approach to understanding the nature of dialogic instruction has been the use of fine-grained discourse analysis to identify and describe how discursive features of high quality

instructional discourse impact student language output (e.g. Klingner & Vaughn, 2000; Sharpe, 2008), patterns of talk (e.g. Chinn et al., 2001), and student participation in pedagogical activities (e.g. Klingner & Vaughn, 2000). These studies draw from a more sociocultural approach to understanding instructional discourse, emphasizing the flow of discussions and patterns of talk opportunities during instruction. The assumption for many of these studies is that engagement in talk with peers and teachers creates an environment that is supportive of academic engagement and each individual's cognitive skills needed to accomplish the curricula. And while this body of work varies in sample size and the explicitness of the coding processes, this research has contributed detailed descriptions of specific features of dialogic talk.

So, though somewhat divergent in theoretical foundations and quite different in their focus on outcome variables, both projects have produced evidence of the importance of participating and receiving academic oral language productions to promote academic language and literacy achievement. Both paths have established several promising features of dialogic instructional talk that likely support students' reading comprehension. Capitalizing on the promise of both can be addressed through a mixed methodology to pedagogical discourse and reading research.

The brief review of extant mixed analytic research into classroom talk suggests several conclusions. Firstly, since talk is a foundational element of sociocultural learning theory and pedagogy, more studies are needed that attempt to not only describe the nature of teacher and student talk in classrooms, but explore the complex, nuanced and highly-situated "discourse-learning nexus" (Murphy et al., 2009, p. 741) that shapes student learning. This lack of discourse analytic studies that include measurements of the impact(s) of instructional talk on student achievement is an important focus for pedagogical research. The use of mixed method

approaches may be a fruitful vein of research to accomplish this important task because this type of research typically attends to some layer of the talk and lesson activity (the level of detail varies) while attempting to explain their relationship to student learning outcomes.

Secondly, there are few studies that looked cross-sectionally at grade level classroom discourse. While there are drawbacks to cross-sectional research into language, at least compared to longitudinal research that can follow individual growth quite precisely, cross-sectional study designs may allow researchers to examine developmental trends of both teaching talk and student language-in-use, and explore how these might be conflictual (Moje, Collazo, Carrillo, & Marx, 2001) or compatible.

Thirdly, very few studies focused on amounts of student talk, which is surprising considering the current interest in inquiry- and discussion-based teaching and the corresponding emphasis in teacher education programs on talk in learning. Those studies that did calculate ratios of teacher to student talk, and the quality of each, draw a rather depressing picture of the limited amount of student talk that is happening in ELA classrooms in the last ten years, with students speaking within authentic discussions on average less than one minute in secondary classes (Gamoran & Nystrand, 1991), and 50 seconds and 15 seconds per class in eighth and ninth grade, respectively (Nystrand et al., 2003). In a sample of fourth-grade classrooms, researchers found that during typical recitational teaching about texts, teachers spoke 53.1% of the total talk (Chinn et al., 2001). Studies emphasizing discourse analysis, for the most part, have provided little in the way of student achievement outcomes, other than to presume that more student talk means better quality learning.

Fourthly, it is heartening to see some studies of bilingual students in mainstream classrooms, lending some ecological validity to this field of research. The notion of ecological

validity stems in part from Bronfenbrenner's (1977) seminal treatise on the importance of researching human development not in a laboratory, but in the environment in which the person interacts. The ecology of human development is the scientific study of "progressive, mutual accommodation…between a growing human organism and the changing … environments in which it lives" (p. 514). Thus, to match the typical schooling contexts in the US, bilingual students' work in both bilingual and English-only school contexts, must be explored.

In sum, this review of research methods into classroom discourse points to a methodological approach that must attend to both language data at the classroom level and individual student literacy outcomes. As Nystrand (2006) has argued in his review of research on the impact of classroom discussion on reading comprehension outcomes, large-scale surveys have been useful in reporting overall discourse patterns and pre- and post-test achievement results using regression or multilevel modeling techniques, but these surveys also limit the notion of classroom discourse to a static macro variable (p. 401). Discourse analytic techniques have captured the fluidity of classroom discourse, but these are not usually linked to student achievement outcomes, nor to contextual factors like demographic variables (ibid.). If classroom discussion affects student learning in indirect ways, a mix of methods that attend to the "epistemic" variables of discussion (Nystrand et al., 2003) as related to achievement outcomes is necessary. Thus there is a need for research to explore the nature of teacher talk during upper elementary literacy instruction and examine its impact on the growth of student reading and vocabulary achievement. This is the express intent of the current research.

Chapter 3 – Methodology of Study

Introduction: Mixed Methodology

In this study, we hypothesize that exposure to high rates of dialogic features of instruction play an important pedagogical role in supporting student reading comprehension achievement in the upper elementary grades. Sociocultural theorists of learning see learning not only socially-situated and deriving from interactions with others (Renshaw, 2004; Vygotsky 1978; Wells, 1999; Wertsch, 1998), but an inherently semiotic process where learners draw on language to make meaning, as "language is the essential condition of knowing" (Halliday, 1993, p. 94). The centrality of language and interaction is captured in the oft-cited definition of reading comprehension as a cognitive process of "simultaneously extracting and constructing meaning through interaction and involvement with written language" (RAND, 2002). If, in the process of using language through interactions with others, one is learning, then students with exposure and the opportunities to engage with peers and teachers through language around texts would also be engaged in acquiring the language needed to comprehend the text. Drawing on insights from both quantitative and qualitative analyses, these intersections between language, interaction, and comprehension are explored. A mixed methodological and philosophical orientation can allow for more nuanced answers to questions about effective features of instructional talk than one approach alone (Bryman, 2006; Creswell & Plano Clark, 2011; Morse, 2010). With a fairly well-established line of research on classroom discourse in both methodological traditions, drawing on both in one study has the potential to clarify the characteristics and outcomes of dialogic interactions around oral and written language.

The use of a mixed analysis can be particularly useful to study the relationship between instruction and outcomes. The overarching purpose of this design is reflected in the research

questions; namely, that a review of the literature on classroom discourse points to the need for a better understanding of its impact on student achievement that a mixed analysis can potentially provide. The study of dialogic instruction necessarily means the analysis of classroom language, but this alone does not fully address educational activities and goals. Both teachers and students are in some way held accountable for the talk in which they are engaged, as it is in this mode that the delivery of content and the assessment of knowledge occur. Therefore, discourse analysis of classroom talk must be linked in some way to student achievement and student learning if it is to be instructive research for educators and policy makers, as well as researchers. With insights from sociocultural and dialogical theories of language learning and statistical methods that account for multiple levels of data, analysts may be more able to handle the complex intersection of language exposure, school pedagogical activities, and student-level outcomes. Since learning in schools is characterized by multiple layers of influences (e.g. individual aspects of learning, peer influences, classroom and teacher variables, etc.), attention to at least some levels of influence provides a more robust picture of the variables affecting each student's learning and developmental outcomes. Research that links these instructional contexts to student learning outcomes has the potential to contribute substantially to improving teaching and learning in school settings.

Mixed research has been typically characterized as a pragmatic philosophical approach to research interested in capitalizing on the "insights provided by qualitative and quantitative research" (Creswell & Plano Clark, 2011; Johnson & Onwuegbuzie, 2004, p. 16). Creswell and Plano Clark (2011) present three other worldviews that have underpinned mixed methods studies, but the foundations of pragmatism are the most relevant to the current study, as it is concerned with the "consequences of (pedagogical) actions" and is oriented to practical

outcomes in typical, "real-world" classrooms (p. 40-41). The current study draws on these potential strengths of a mixed methodology to explore the relationship between dialogic instruction and student reading achievement.

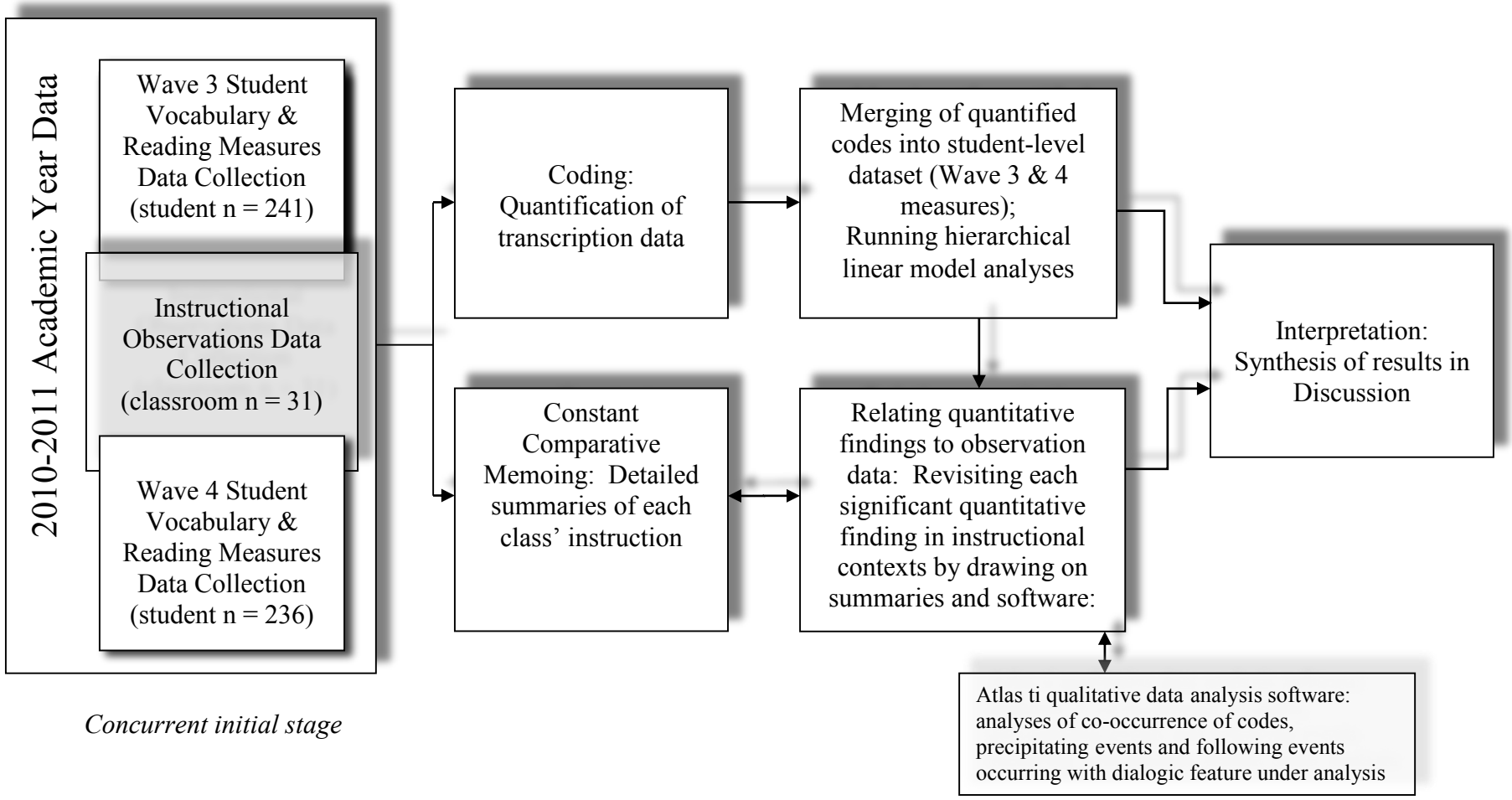
Convergent Mixed Methods Research Design

The "mixing" in mixed methods refers not only to the blending of quantitative and qualitative analytic approaches, but the determination of when, where, and how the strands are related and integrated. In the current study, the overarching research question, how the quantity and quality of dialogic instruction affect student reading comprehension, addresses the integrated nature of the mixed methodological plan. To answer this question and the quantitative and qualitative sub-questions, the study follows a convergent design with features of explanatory sequential design in the analyses (Creswell & Plano Clark, 2011). Student vocabulary and reading measures (quantitative data) were collected concurrently with the classroom observations, and then analyzed separately and sequentially before converging the results during the last phase of data interpretation (Tashakkori & Creswell, 2007). Research questions 1 and 1a focus on the relationships between student achievement scores and the instruction students were exposed to during the school year. To answer these questions, quantitative methods were used that depended first on the qualitative coding of instructional discourse (Johnson & Onwuegbuzie, 2004). Research questions 2 and 2a address characteristics of both qualitative and quantitative research. After running linear models to explore the relationships between dialogic features of instruction and students' reading scores, statistically significant findings were then explored by analyzing their patterns and variations in the language arts lessons. This sequential design converges at three points in the analysis: 1) when transforming qualitatively-based coding of instructional talk to use in the quantitative modeling; 2) when integrating the significant

quantitative findings to guide the second stage of qualitative analysis; and 3) by summarizing both qualitative and quantitative results as to what extent they relate and explain each other (Creswell & Plano Clark, 2011). This analytic design also draws on principles of explanatory mixed method design in that a key purpose of the qualitative analysis is to explain the significant findings from the quantitative analysis, and using the quantitative results to make decisions about the direction of the qualitative analysis (Creswell & Plano Clark, 2011).

This study can be characterized broadly as a QUANT → qual study that is organized by research questions from different paradigms and driven by questions about the relationship between instructional characteristics and student outcomes. The analysis is mainly done sequentially (hence the arrow) but both components interface at various points in the analysis (Morse, 2010). This attention to the strengths of both paradigms is one advantage of mixed methods research (Teddlie & Tashakkori, 2003). It is used here to also allow for better triangulation of the classroom data: the multilevel modeling of student achievement scores nested in classrooms allows for the validation of theories of dialogic instruction, and relies on discourse data to both identify contextual factors of elementary literacy instruction, and describe the complex nature of instruction moves. Thus, these instructional and achievement data have both global and specific aspects that warrant attention, and a mixed analysis can address such aspects concurrently, as well as provide more explanatory analysis of correlated findings. This mixing can also provide a more comprehensive account of reading instruction, and explain any unexpected results from the QUANT component of the analyses. By attempting to illustrate the quantitative findings with qualitative analyses, this direction can also enhance the utility of findings for instructional implications (Creswell & Plano Clark, 2011). Figure 3.1 visually summarizes the mixing of the data and analyses in the current study.

Figure 3.1 Overview of the Study's Convergent Mixed Method Design



Research Design

Context of the present study: The CLAVES study.

The data in this study are part of a parent study on the literacy development of English monolingual and Spanish-English bilingual students in grades two through five. The CLAVES[1] project (Comprehension, Linguistic Awareness and Vocabulary in English and Spanish) followed students over two years of schooling to track their reading comprehension, their breadth and depth of vocabulary development, and their literacy instruction. The students, in six elementary schools (three in Massachusetts and three in Maryland), were given a battery of language assessments in the fall and spring of the 2009-2010 and again in 2010-2011, and the literacy/language arts classes of consenting teachers were observed three times over the 2010-2011 academic year. Data for the current study are drawn from the 2010-2011 school year with students in the observed classrooms in grades three, four, and five.

Maryland sites.

The three schools in Maryland were K-5 elementary schools in the same public administrative district. Two were in an urban area and one was located in a rural area about a half hour drive from the other two. One urban school was surrounded by main roads, a shopping area, and a new upscale condo complex; the other was nested in a working class neighborhood bordering a large utility storage area. The rural school was surrounded by working class homes on ¼ acre lots in an area home to many casino businesses. All three school buildings were too small to service the current number of students, so at all three schools, the fifth-grade classes

[1] The research reported here was funded by a grant from the Institute of Education Sciences, U.S. Department of Education, to the University of Maryland (No. R305A090152; Silverman (PI), Proctor (co-PI), and Harring (co-PI). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

were housed in modular classrooms next to each school building. All schools were English-medium, public elementary schools.

The student populations at the Maryland sites were majority African American or Latino with small numbers of Asian and White students. The largest percentage of White students was in the rural school (see Table 3.1).

Table 3.1 *Percentages of Students by Race, Language and SES Status in both Districts*

	White (%)	Latino	Black	Asian	Bilingual	ELL/LEP	FARMS
MA District	54.4	27.6	8.6	6.8	35.9	9.6	37.0
School 1	60.1	22.2	7.0	8.6	36.1	55.2	35.4
School 2	54.9	25.2	9.1	7.6	30.6	5.0	37.2
School 3	32.4	57.6	5.6	1.9	54.2	24.6	34.3
MD District	62.0	8.6	21	3.4		5.7	33.9
School 4	23.0	16.0	45.0	11.0			
School 5	3.1	41.0	48.3	0.0			
School 6	2.8	59.2	32.6	2.5			

Note. Unavailable data left blank; MA = Massachusetts; MD = Maryland; ELL = English Language Learners: students qualify to receive SEI support (parents can opt-out; term used by MA district); LEP = Limited English Proficient: students enrolled in ESOL support classes (term used by MD district); FARMS = Free and Reduced Meal Students, a proxy for socioeconomic status

Massachusetts sites.

The three schools in Massachusetts were K-5 elementary schools in a semi-urban district. Each school was only five years old, newly built or renovated and relatively spacious, though one had reached capacity during the 2010-2011 academic year. All three bordered residential areas, and two of the three were centrally located near shops, the public library and small businesses. The third was in a middle-class neighborhood of wooded lots about 10 minutes from the centre of town. Like the Maryland schools, these three were English-medium public schools.

The student population in the Massachusetts district was majority White with a large percentage of Latino/a students, and a smaller percentage of African American and Asian students. The two more centrally-located schools reflected the city's shifting demographics more than school 1 (10 minutes from the town centre). School 3 was majority Latino/a (see Table 3.1).

Data Samples for Present Study.

There were two samples in this multilevel study: teachers, and students clustered by teacher.

Teacher participants.

The present study analyzed data from 31 teachers. Two of the 33 teachers in the parent CLAVES study were not included in this study because they did not meet an acceptable number of students participating in the assessment measures over the 2010-2011 school year (one student only). All teachers that were observed for the current study were the primary instructors in their classrooms, and most were certified in elementary education. All taught in grades three, four, and five. Four teachers were certified in reading, and one in teaching ESOL. Of the 31 teachers, 13% were male, 81% were Caucasian, 3% where African American, 3% chose 'other' to describe their race/ethnicity, and four teachers did not report. Teachers had taught between one and 35 years at the K-6 level (M = 7.8 years, SD = 7.6 years). Two teachers did not report their years of experience. Most held masters' degrees or 30 credits beyond a masters' degree (72%) and were certified in elementary education (69%, with 31% of teachers not reporting).

Other relevant professional experience in this sample of teachers, including percentages of teachers' degree type, years taught, completed PD courses addressing LM learners, and bilingualism, are presented in Table 3.2.

Table 3. 2 *Characteristics of Teacher Sample (%)*

Characteristic	%	N
Degree		
BA	28	8
MA	69	20
MA+30	3	1
Years taught K6		
1-5	55	16
6-15	27	8
16-30+	17	5
Number of PD courses on LM students		
0	18	5
1-2	57	16
3+	25	7
Bilingual	10	3

Note: Self-reported data, based on voluntary teacher survey; two teachers missing 'degree' and 'years taught' data, 3 teachers missing 'PD courses' data

Twelve teachers participated in the classroom observation data collection in the three Massachusetts schools, and 19 observed teachers in Maryland. The majority of teachers were observed three times over the 2010-2011 academic year. There were five teachers who were missing one observation: one teacher declined a third observation, the audio recordings of two observations in two classrooms were corrupted, and due to scheduling conflicts, we could not get literacy/language arts teaching observations of two other teachers. All observed classes were mainstream classrooms with the exception of one fourth-grade Sheltered English Instruction (SEI) classroom for newly immigrated English learners at the Massachusetts site. Third- and fourth-grade classrooms were observed in all six schools, but only four of the six schools provided fifth-grade classroom observations to this sample. Table 3.3 summarizes the total number of classrooms observed at each grade level at each site.

Table 3.3 *Number of Teachers by School and Grade Level*

	3rd Grade	4th Grade	5th Grade	**Total**
Massachusetts				
School 1	2	2*	1	5
School 2	1	1	1	3
School 3	2	2	0	4
Totals	**5**	**5**	**2**	**12**
Maryland				
School 4	2	2	0	4
School 5	3	2	3	8
School 6	2	2	3	7
Totals	**7**	**6**	**6**	**19**
Total Classrooms	***12***	***11***	***8***	***31***

**One Sheltered English Instruction (SEI) class of newcomer Latino/a students*

Student participants.

The second sample in this multilevel dataset was the student sample; a subset of the 2010-2011 CLAVES cohort, who were students who attended the classes in which we observed language arts instruction in grades 3, 4 and 5 (student $n = 236$). Teachers were observed to have between 12 and 26 total students in their classes. A subsample of students from each classroom was selected to participate in the CLAVES study. The observed teachers had between two and 12 ($M = 7.6$) students who participated in the reading and vocabulary assessments, and these 236 students are the focus of the current study.

Demographic data from parent surveys and school data showed that monolingual English-speakers at the Massachusetts site were mainly Caucasian, while monolingual English-speakers in the Maryland schools were predominantly African American. By design, all language-minority students in both districts were Latino/a Spanish speakers. There were 138 students in the sample that listed English as their first language (58% of the sample), and 103 students that were Spanish-English bilinguals (44%). For the study, we considered students to be

bilingual if their parents reported Spanish spoken in the home on the parent surveys. In the sample, 98 students spoke Spanish as their first language (42%), and of these students, 69 were classified as English learners that qualified for ESOL or SEI support (Maryland and Massachusetts districts, respectively). The Maryland schools had a larger proportion of L1 Spanish speakers, but the Massachusetts sample had a larger proportion of bilinguals that were identified as English learners compared to the Maryland cohort of bilingual students. These language data are presented in Table 3.4.

Table 3.4 Student Language Status by Site

	Maryland Sample		Massachusetts Sample	
Characteristic	N	%	N	%
L1 English	90	54	48	70
L1 Spanish	77	46	21	30
Bilingual Non-EL	57*	34	9**	13
Bilingual EL	20*	12	17**	25
Total Students	167		69	

Note. Student $n = 236$
* 4 students missing data; ** 8 students missing data

When comparing this sample to both districts' demographics, bilingual and EL students were oversampled for this study. Limitations of this decision will be discussed in Chapter 6, however, for this study, this oversampling reflects the importance of ecological validity for research on instruction by addressing the underrepresentation of bilingual students in such research; the Nystrand studies (Applebee et al., 2003; Gamoran & Nystrand, 1991; Nystrand & Gamoran, 1991) tended to be monolingual populations for instance.

The number of students clustered with teachers at each school and grade level is summarized in Table 3.5.

Table 3.5 *Number of students by Teacher, by Grade, and by School*

Grade	Teacher ID	Site	School	Student n
3				
	4	1	1	5
	6	1	1	6
	16	1	2	6
	22	1	3	7
	23	1	3	7
	50	0	4	2
	54	0	4	6
	62	0	5	11
	63	0	5	12
	64	0	5	10
	79	0	6	12
	84	0	6	8
Sub Totals	***12***			***92***
4				
	9	1	1	6
	10	1	1	4
	18	1	2	5
	25	1	3	4
	26	1	3	8
	53	0	4	3
	65	0	5	11
	66	0	5	11
	70	0	4	9
	73	0	6	10
	75	0	6	9
Sub Totals	***11***			***80***
5				
	27	1	1	5
	30	1	2	6
	81	0	5	8
	82	0	5	8
	83	0	5	10
	85	0	6	9
	86	0	6	10
	88	0	6	8
Sub Totals	***8***			***64***
Grand Totals	**31**			**236**

Note. Site: MD = 0, MA = 1; Student n is derived from those students that completed both waves of the outcome measure of interest in this study, the Woodcock Munoz Language Scales- Revised Passage Comprehension (WMLS-PC) subtest.

Although we intended to have a minimum of five students tested on all measures per teacher, student attrition and absences during the 2010-2011 school year led to some missing scores on Fall (Wave 3) and Spring (Wave 4) measures. Students with missing scores on these measures were dropped from the analysis to make sure that the students exposed to their teacher's instruction over the academic year were also the same that had Fall and Spring scores. The total number of students in Table 3.5 reflects the number of students that completed the Fall and Spring reading comprehension outcome measure for the study. In the final sample, only four classrooms had less than five students with complete Spring and Fall reading comprehension measures used in the final models (Teachers 50, 10, 25, and 53 had 2, 4, 4, and 3 students respectively).

Data Collection Procedures.

At the beginning of the 2009-2010 school year, student participants in grades two through four were solicited through letters to their parents/guardians to participate in the CLAVES study. Of the pool of students whose parents consented, we randomly selected monolingual English and bilingual Spanish-English students. Given the relative homogeneity of the monolingual populations at each site, race/ethnicity was a selection criterion for monolingual students as a means to promote as much racial and ethnic variability as possible among the monolingual sample. At the beginning of the 2010-2011 school year, we solicited more students to participate in the larger study, this time in grades three through five. This new cohort was added to the first year cohort so that the CLAVES study had outcome measures on a cross-sectional sample of students in grades two through five. Students were assessed on a variety of standardized

vocabulary, reading, and writing measures. They were taken out of their classes for about two hours, once in the Fall and once again in the Spring before each state's testing season began, to be tested in small group and one-to-one settings. Teachers decided on the times students could be taken from class to complete the testing in order to minimize disruption. Testers were trained graduate students from both research sites. Participating students were given small "prizes" (erasers, pencils, small toys) as thanks for participating in each of the testing sessions. Data for the present study were from the second year of the parent study, and are thus comprised of English monolingual and Spanish-English bilingual Spanish-English speaking students in grades three, four, and five.

Teachers were solicited to participate in the teaching observations in September, 2010. Of the pool of consenting teachers, we purposively selected teachers at each grade level with at least five students in their classrooms who were participating in the CLAVES study that year. All observed instruction was with teachers in mainstream classrooms with the exception of one teacher at School 1 who taught a fourth-grade SEI class for English learners (ELs) who had immigrated to the US within the last two years, and were assessed as needing English language support beyond what was provided in mainstream classrooms. Her students remained with her for their content-area instruction daily.

Classroom Observations.

The 31 teachers were observed three times over the course of the 2010-2011 year, once in Dec/Jan, again in Jan/Feb, and for a third time in Feb/Mar. In similar research, researchers have argued that observations of instruction collected at three time points are acceptable to capture a teachers instructional tendencies (Connor et al., 2011; Connor, Morrison, & Petrella, 2004) because instruction has been shown to be quite stable across time (Smolkowski & Gunn, 2012).

All observations occurred during literacy or language arts instructional time. Teachers at the six schools used various terms for this time ("literacy time," "workshop," "language arts," "ELA," "reading time"), but they all engaged in activities that typified primary school language arts instruction: learning tasks such as reading aloud, talking about readings in small or whole class groupings, modeling and applying reading strategies, writing on worksheets or in writer's workshop formats, vocabulary instruction linked with text, grammar and other form-focused instruction that teach the "rules" of the English language and writing, and reading comprehension and vocabulary assessments.

For simplicity, the term *language arts* will be used, except in circumstances when a more specific term is needed to clarify the classroom activity. Sustained silent reading was not observed since this activity is usually not characterized by teacher instruction. These language arts instructional blocks were observed between November and March to capture instruction after students and teachers had settled into instructional routines and before each state's testing season began in the spring. Language arts instruction was targeted for observations to document the bulk of literacy instruction that students experienced throughout the school day. While these teachers undoubtedly provided opportunities for students to use high-level literacy when teaching Math, Science and Social Studies as well, like other studies have shown, these lessons tended to focus on the conceptual knowledge of their respective curricula, and not on explicit reading or writing skills for the specific content area (Moje, 2008). And while "the goals of English (ELA) instruction have always been diverse, involving mastery of virtually every activity connected with the use of language" (Applebee et al., 2003, p. 687), pilot observations found that this sample of teachers explicitly taught reading and writing skills mostly during their language arts time. Only once did one teacher use her "literacy" time to teach science content

through small reading groups with leveled readers. In this lesson, the focus was as much about accomplishing the reading as it was about understanding the concepts of electricity: students were asked to practice certain reading strategies while building their understanding of topics in electricity addressed by the texts.

Teachers chose the times of each observation and were encouraged to teach as they normally did. Observers were trained research assistants associated with either site, each with a background in teaching, and were trained to take field notes using video from three pilot CLAVES classroom observations. Observers audio recorded as much of the classroom talk as possible by placing digital recorders in a central location in each classroom and took field notes on laptops to capture non-verbal instruction like teacher movements around the classroom, materials used during instruction, gestures, classroom resources evident in the class, some one-on-one student-teacher interactions, and student participation (hands raised to teacher questions, time on task, qualitative judgments on how many students were on task). Observers organized their notes into three minute intervals to better merge the audio with these field notes.

The audio was professionally transcribed by two services. The transcriptionists were asked to transcribe all teacher and student talk that was audible, and mark where the talk was unintelligible. Transcription necessarily demands the analyst to make decisions of relevance, which are ultimately theoretical judgments since they are based on the researcher's theories of how language, situations, and interactions work (Gee, 2005, p. 106). We took a pragmatic view of the transcription for the CLAVES study, focusing on the lesson flow and content of the lessons through the talk and the field notes. And since the present study is not focused on the micro interactions of speech utterances like prosody, inflections, and pauses ubiquitous in most speech, nor on individual student contributions, these details were not noted in the transcriptions.

Transcriptions thus documented the audible talk from the teacher, any assistants in the classroom, and students for the duration of the language arts lessons, keeping the identity of the students anonymous. Field notes were merged with these transcriptions to form the classroom observation data set.

Each observation varied in length, from one lesson that was only 17 minutes because of a fire drill, to some observations that were almost two hours long that included two blocks of literacy instruction. There were 88 transcripts analyzed from the 31 teachers observed. The average length of the observations was 55.35 minutes (SD = 16.82 minutes), with a range of 17 minutes to 115 minutes. These lengthy observations generated multiple recordings in a single wave of observations. In cases where there were multiple transcripts for one teacher on a single day, I strategically chose to code the observation(s) that had the most students in the class during the lesson observed. If these numbers were not available (for example, an observer forgot to note the number of students in one class during the observation), I chose to code the observation that demonstrated the most amount of language arts and literacy class work of the available observations. For instance, in some classes in the Maryland site, some literacy lessons were called "workshop" lessons, which ranged from catch-up work in a range of content areas, to small group reading activities and vocabulary spelling tests. If these were the observed activities, I chose to code the small group reading activities since they are most closely related to the language arts curriculum and to the goal of reading instruction. This purposeful selection of portions of some teachers' ELA instruction was applicable to only a few transcripts. The varying length of time of the observations was controlled for by converting raw counts of codes to code rate per minute of instruction. These rates were used as covariates in the final models.

Student Literacy Measures.

The sample's 236 students were tested on various literacy measures in the Fall and Spring of the 2010-2011 school year.

Reading Comprehension. To assess students' abilities to comprehend sentence- and passage-level texts, the Woodcock Muñoz Language Scales, Revised, Passage Comprehension (WMLS-PC; Woodcock, Munoz-Sandoval, Reuf, & Alvarado, 2005) subtest was administered twice over the year. The WMLS-PC is an untimed test administered by a tester with one student at a time. Students read a series of progressively longer and more challenging sentences and paragraphs, either silently or aloud, and were asked to provide a word to fill a cloze blank within the text. Responses were scored as correct or incorrect according to the context of the text. Reliability of this subtest for children between 8 and 11 years old is.81 to .91 (Woodcock et al., 2005). Students' Fall (Wave 3) and Spring (Wave 4), age-equivalent standardized scores were used for this analysis. Standard scores were used as outcomes to allow us to generalize findings to the average student in the average classroom.

Vocabulary Measures. Vocabulary comprehension is a well-established correlate to reading comprehension (Anderson & Freebody, 1981; Duke & Carlisle, 2011; Joshi, 2005; Snow, Burns, & Griffin, 1998). To capture the breadth and depth of students' vocabulary knowledge, various assessments were administered. The following measures were used in the final models to control for students' initial vocabulary knowledge.

As word recognition and expressive vocabulary breadth are well-established correlates of reading comprehension, it was important to account for students' skills in these areas (Geva, 2006; Hoover & Gough, 1990; Snow et al., 1998). The WMLS-R Letter Word recognition subtest (WMLS-LW; Woodcock et al., 2005) was administered to students in a one-to-one

setting. The WMLS-LW subtest measures increasingly difficult word recognition. Ceiling is reached when the student reads aloud six consecutive items incorrectly. Internal reliability of this test is .98 for children aged 8 year, and .96 for 11-year-olds (Woodcock et al., 2005). The WMLS-R Picture Vocabulary (WMLS-PV) subtest was administered to capture students' expressive vocabulary breadth. Students name pictured items of increasing difficulty or obscurity and reach ceiling with 6 consecutive wrong answers. Internal reliability ranges from .90 to .92 (8-year olds and 11-year olds, respectively; Woodcock et al., 2005). Standardized scores, centered on each student's classroom mean, on both tests from the Fall 2010 (Wave 3) test administration were used in the final models.

Vocabulary depth has been shown to contribute to students' reading comprehension achievement (Proctor et al., 2012). Student semantic, syntactic, and morphological knowledge at the beginning of the year was included in the final models to control for students' initial vocabulary knowledge.

The Clinical Evaluation of Language Fundamentals Word Classes 2 subtest (CELF-WC; Semel, Wiig, & Secord, 2003) was used to measure students' semantic understanding. Individual students were read aloud four words and asked to indicate which two of the four were semantically related ("*Which words go together: teacher, student, cake, street?*"). Ceiling was reached with 5 consecutive missed items. Internal consistency (coefficient α) for students 7.0 – 13.11 years old is .73- .84 (Semel et al., 2003).

The CELF Formulated Sentences (CELF-FS) subtest was used to measure students' syntactic and semantic knowledge. Students were shown a picture and given a target word to use in a sentence describing the picture. Testing was discontinued after 5 incorrect responses,

whether grammatical or semantic mistakes. Internal consistency (coefficient α) for students aged 7.0 – 13.11 years is .75-.82 (Semel et al., 2003).

Finally, morphological knowledge was measured as the third component of vocabulary depth using the Extract the Base test (EB; Anglin, Miller, & Wakefield, 1993; August et al., 2001; Carlisle, 1988). The EB test was another individually administered test, requiring the student to use the base word from a derived word (e.g. *farm* from *farmer*) contextualized in a sentence (e.g. *My uncle works on a* _____). Students followed each sentence on their own papers while the tester read the words and sentences aloud, and then students wrote the appropriate word in the blank. August et al. (2001) report Rasch-based reliability at .98. Classroom mean centered raw scores were used for each of these three tests above measuring students' semantic, syntactic, and morphological knowledge in the Fall.

Analytic Plan

This study follows a convergent parallel mixed method design with features of explanatory sequential design in the analysis (Creswell & Plano Clark, 2011), as summarized in Figure 3.1, above. After the concurrent data collection of the two sample datasets (student-level literacy achievement and classroom-level lesson transcripts), transcripts were analyzed through coding with both an *a priori* coding scheme at the utterance level of classroom discourse, and through emergent coding at a more global level of classroom instruction (description follows). These codes were then quantified and merged with the student assessment data to form a two-level dataset. Hierarchical linear modeling was then used to explore the relationship between features of instruction and student reading achievement scores. Statistically significant findings from the models were then further explored by qualitatively analyzing them back in the transcription dataset to provide a more complete understanding of how these significant utterance

level features function during lessons to support students' reading achievement. These findings were then considered together to discuss how the two types of data were related (Creswell & Plano Clark, 2011). This analytic plan follows a sequential design by transforming results from the qualitative coding of observational data into the quantitative analyses to inform a qualitative analysis of these significant findings to produce a data-driven explanation of the efficacy of certain dialogic features of instruction. This analytic framework sought to build on the strengths of both correlational and discourse analytic studies of classroom talk to provide a more comprehensive account of instructional talk through the use of qualitative data to illustrate significant quantitative findings as well as to improve the usefulness for practitioners (Bryman, 2006). Converging both types of data also allows for better triangulation of the data so that data are better corroborated (Bryman, 2006; Creswell & Plano Clark, 2011). Specific details of both the qualitative and quantitative analyses are discussed in the following sections.

Qualitative analytic approach.

There were two stages of the analyses that used qualitative methods. The first was to code the transcripts of classroom instruction using an *a priori* coding scheme drawn from findings in the classroom discourse literature. The second qualitative analysis was done after the quantitative analysis, taking the significant findings from the HLM models and analyzing them within the contexts of their lessons.

First stage of qualitative analysis: Procedures for coding observational data.

When researchers take a broad approach to analyzing conversation, a coding manual is typically developed to specify units of analysis, and statistical methods can be used to contrast discourse patterns (Sawyer, 2006, p. 200). It has been argued that "coding qualitative data differs from quantitative analysis, for we are not merely counting. Rather, we are attaching codes as a

way of identifying and reordering data, allowing the data to be thought about in new and different ways" (Coffey & Atkinson, 1996, p. 29). In this study, codes were used to identify instructional data and reorder concepts of instructional talk in the empirical literature, and then converted to frequencies that were counted in the later quantitative analyses. A coding scheme (Appendix A) was developed from the empirical literature to identify features of dialogic instruction in order to answer RQs 1 and 2. Previous discourse analytic and correlational work was drawn upon here to establish this coding scheme so that the analysis could move beyond minutely detailed descriptions of a small sample of classrooms.

Glaser and Strauss (1967) have noted that using *a priori* categories to analyze qualitative data may shift the analytic process toward a less generative process and instead toward simple data selection (p. 37). However, the express purpose of the current study was to identify and explore dialogic indicators based on past empirical evidence and a pragmatic view of analysis. Furthermore, the *a priori* coding categories have "conceptual congruence," which is to say that they make sense together (Merriam, 1998, p. 184), because of the long line of classroom research studies that have documented instructional moves (e.g. Applebee et al., 2003; Nystrand & Gamoran, 1991; Nystrand et al., 2003; O'Connor & Michaels, 1993) and patterns of pedagogical talk (e.g. Cazden, 1988; Mehan, 1979; Sinclair & Coulthard, 1975). For instance, it is clear from these transcripts that these classrooms reflect similar styles of teaching documented in other classroom observation work that has focused on instructional discourse (Cazden, 1988; Mehan, 1979). Teacher questions tended to drive the talk, and evaluations tended to follow student contributions. Furthermore, codes were developed in a more generative way to capture features of instruction that were not talk but may mitigate the instructional talk that occurred in these classrooms. Thus, as with all qualitative analyses, inductive and deductive processes were used;

not only by explicitly using both an *a priori* coding scheme and more grounded coding for different aspects of the transcripts, but also when refining the emergent codes using the constant comparative method. As Merriam (1998) explains, even in grounded theory, there is a constant shifting between these two modes of thinking when newly-derived codes are tested against the data to see if they are supported by the data (p. 192). Thus, this deductive process was based on inductive code generation. Therefore, while this study recognizes Glaser and Strauss' (1967) important critique, an *a priori* coding scheme can be justified with the study's RQs and is complemented by some emergent coding incorporated in the final coding scheme.

The coding scheme was organized at two levels of analysis: 1) global, or episodic, features of lessons, and 2) utterance level features of dialogic instructional talk. The global features of instruction were coded to indicate the context in which the dialogic indicators were used. At the utterance level, features of instructional talk were coded for discreet instructional moves made by all participants involved in the lessons.

Global features of instruction.

To get at the more general patterns of instruction for these grade levels, inductive, emergent coding was used to document who managed the discourse, what materials were used, and what topics were addressed during each lesson. An instructional episode is a coherent activity centering around a particular objective or purpose (Nystrand et al. 2003), "bounded sequences of actions which are somehow coherent internally" (Linnell, 1998, p. xiv). A new episode was marked when the teacher addressed a new objective for the lesson; for instance, a shift in activity from a read-aloud to writing about the story. Instructional episodes in the 31 classes observed ranged from about 5 minutes to 60 minutes.

Each instructional episode was coded for 1) teacher- or student-management of the learning activities, 2) instructional materials, 3) genres, or patterns, of instructional talk, and 4) the field, or topic, studied during the instructional episode. Three of the four global codes were developed using the constant comparison method after reading through each transcript. Each instructional episode of each lesson was identified by teacher procedural talk (e.g. "Put your books away and take out your writing journals") or by the field notes. The four global codes were applied to each episode based directly on the evidence from the instructional talk and/or field notes. Teacher or student management of the learning activity was coded by all three coders and a reliability coefficient was established (described below). The other three facets of the lessons, instructional materials, talk genres, and the language arts topics addressed, were coded inductively and refined by the first coder using a constant comparative method of coding. Since these codes were open-ended (for example, there was no limit to the possible materials used or language arts topics addressed), these were first simply identified (e.g. *Sign of the Beaver* novel) and then refined to more general categories by checking for similar groupings with each successive transcript (e.g. unabridged literature). The patterns and regularities in the 31 teachers' use of materials, genres of discourse, and topics covered became the descriptive codes for each instructional episode finalized in the coding scheme (Merriam, 1998, p. 181; see Appendix A).

The teacher- or student-managed activity code was applied to each instructional episode to identify who was responsible for forwarding the learning activity: teachers or students (Connor et al., 2007). If students were writing individual responses to a reading, this would be coded as a student-managed episode, as the main emphasis of this activity was on students managing their own work; if the lesson involved a whole-class, text-based talk with the teacher posing questions, this would be a teacher-managed episode since the teacher is responsible for

moving the lesson along through teacher directed talk. The second global code, instructional materials, is, like pedagogical talk, a semiotic tool for comprehension, a key concept in activity theory (Wells, 1999) and distributed cognition (Pea, 1993). Instructional materials were resources like basal readers, unabridged literature, and grammar worksheets. The coding of talk genres and the fields (topics) of lessons draw from Systemic Functional Linguistics (SFL) theory which foregrounds the culturally specific patterns of pedagogical talk and curriculum to understand language choices made during instruction (Christie, 2002). A talk genre has been defined by Christie (2002) as a "staged purposive activity undertaken to accomplish some goal" (p. 21), and in the current data, these were discourse activities like read alouds, whole-group text-based talks, and individual seat work. The fields of lessons were language arts curricular foci like text comprehension, reading strategies, explicit vocabulary study, and test preparation.

Each of these four global features of lessons was quantified by tallying each code for each teacher, and frequencies were calculated across the total number of transcripts (Chapter 5, Tables 5.1, 5.2; Appendix B).

Utterance-level features of instruction.

At the utterance level of transcripts, features of instructional talk were coded and quantified. Utterances were defined as the smallest segment of intact talk by a student or teacher that conveys an idea: "Nice work" or "Why do you think that?" This is similar to Crookes (1990) definition as a unit of speech expressed by similar intonation bounded by pauses. Only pedagogical talk that addressed and/or forwarded the content of the lesson, and not procedural, managerial, or rhetorical talk, was coded for the following talk moves.

Authenticity of teacher questions. Students tend not to ask inauthentic questions when on task (Nystrand & Gamoran, 1991) so this was a particular measure for teacher questions only. However, measuring authenticity implies an understanding of teachers' intentions, data that was

not collected for this study. Per Nystrand and his colleagues' work (1997), three levels of teacher questions were considered in this coding scheme.

1) *Authentic questions* (AQs) were those that did not have a prespecified answer that the teacher was seeking and do not show uptake of previously contributed ideas (Applebee et al., 2003, p. 700), as shown through her response to students' responses.
2) *Quasi-authentic questions* (QQs) were those that offered students "some latitude in answering" but have a finite range of possibilities (Gamoran & Nystrand, 1991, p. 288). Synonyms for a vocabulary word were a common quasi-authentic question in these data.
3) *Test questions* (TQs) were those questions posed when the teacher and students were presumed to know the answer, and the question had only one possible correct answer (Boyd & Rubin, 2006, p. 152; Cazden, 1988; Chinn et al., 2001). These TQs suggest a lack of dialogic instruction, a lack of reciprocity to the pedagogical talk.

A fourth question type drew on more recent studies of teacher questioning.

4) *Uptake questions* (UP) were teachers questions that extend a previous student utterance (Boyd & Rubin, 2006), by connecting to and 'picking up on' previously stated ideas (McNeill & Pimentel, 2010). Uptake was an indicator of the reciprocity of talk on a topic, marked by teacher questions that incorporated a previous response. Students usually exhibit uptake in their questions if they are attending to the discussion, while teacher uptake varies (Nystrand & Gamoran, 1991, p. 273).

Since talk used to set-up and managed learning activities was not coded, procedural, managerial and rhetorical questions were not included in the analyses.

Each question code was tallied from each of the classroom's observation transcripts, and weighted means were calculated in three different ways. Totals were divided by the number of minutes of each transcript, and averaged across each teacher's three classroom observations (per minute classroom mean). Rate of questions per word was also calculated by dividing the total number of questions by total number of words in each transcript. I also calculated the average percentages of each type of question to the overall number of questions each teacher posed. For the final analyses, only the question rates per minute were used as predictors in the models presented in Chapter 4. These rates were centered on classroom means (discussed below).

High- and low-quality questions. A second group of codes was applied to utterance-level data to identify the cognitive quality of questions asked of students. As students in elementary school transition from learning to read to reading to learn, they are increasingly required to use skills of analysis, evaluation and synthesis to learn from content-area texts. For this reason, the coding of what cognitive skills each question elicits from students may be helpful in understanding what kinds of cognitive skills teachers are attending to in their lessons, and how these may relate to students' reading skills. Each question was coded in terms of what kind of thinking they elicited, based on the student response and the field and genre of each instructional episode.

1) *Low-level questions* (QULow) were considered recitation and recall questions where the answers to these questions could be found in the focal text or in preceding talk during the lesson. Nystrand et al. (1997) have called recitation questions "what happened questions" that ask students simply to retell what they have come to know (p. 198). I included fill-in-the-blank questions and definitional questions under this code when it was clear that the answer was presumed to be known.

2) *High-level questions* (QUHigh) were those that asked students to *generalize*, displaying their knowledge of what happens when a phenomenon occurs (Nystrand et al., 1997, p. 198), *analyze* why something happened, *speculate* or predict based on what they already know or infer, and *evaluate* a text or character motives, for example, to display logical reasoning. During coding, coders found that AQ and QQ questions could be either high or low cognitive quality (e.g. such as a review question with a few answers that students just had to memorize), but TQ questions were always of low cognitive quality. Thus, the TQ code along with the QULow code was somewhat redundant, indicating a high correlation of $r = .83$ ($p < .001$; see Table 4.9).

Like the question type codes, these question quality codes were also totaled and converted to rates based on minutes and total words spoken in each transcript, and then averaged across each teacher's transcripts to create a weighted mean for each classroom. Rates per minute were used in the final models because they account for the differing observation lengths. These variables were centered on the classroom means. Percentages of high and low quality questions to each question type (AQ, QQ, TQ, and UP) were also calculated but not used in the current analysis.

Evaluation of student responses. A second group of talk moves coded in these data also applied to teacher utterances only, as students do not typically evaluate their teachers' questions or explanations aloud. These were:

1) *High-level evaluations* (TEvalH) were codes indicating when the teacher acknowledged that a student had contributed something new that changed or modified the topic or direction of discussion (Nystrand & Gamoran, 1991; Nystrand et al., 1997) by explicitly commenting on the student's response, or implicitly addressing

the contribution by using the student's contribution to further an explanation, or further the development of that topic. This can be done by "revoicing" (O'Connor & Michaels, 1993), or "recontextualizing" (Cazden, 1988) a student 's contribution to align it with the academic task, including the facilitation of students alignment with "intellectual socialization" (O'Connor & Michaels, 1993, p. 319). Teachers can also "recast" students' contributions to model more academic language (Sharpe, 2008).

2) *Low-level evaluations* (TEvalL) coded for teacher responses that indicated no or a shallow evaluation of a student's response, a response that did not forward the lesson topic, such as "Good. That's right" and moving on to another question or student contribution. Though these low-level evaluations do not seem to engage students in dialogue, there is some suggestion that they may be indicative of more student discussion (Nystrand et al., 2003). Teachers may repeat a student's utterance, or respond "good" as a conversation placeholder so s/he retains some management of the talk, while also allowing the teacher to get out of the way of student talk.

Like teacher questions, teacher evaluations were tallied per transcript, divided by the number of minutes and the number of words spoken in each transcript, and then averaged across the three observations per teacher. Average percentages of both types of evaluations to each transcript's total number of evaluations were also calculated. Teacher evaluation rates per minute, centered on the classroom means, were used in the modeling discussed later in this chapter.

Teacher or student explanations. The teacher explanation code was applied to teachers' utterances that forward the lesson by explaining a concept or topic without student input and without posing questions. There is evidence that extended explanations are supportive of young

students' language development (Aukrust, 2007), but this could also be considered non-dialogic, lecture-type talk if extensive and marked by few student contributions. Teacher explanations (TExp) were coded when teacher talk was not procedural or managerial, and the utterance was more than two lines of transcript. Student explanations (StExp) were considered utterances of more than two lines of transcript that expressed a coherent idea. This code functioned as a proxy for an increased level of student talk, which has been shown to be associated with reasoning and high-level thinking (Soter et al., 2008). Elaborations in science and math classrooms have been shown to increase student talk during lessons and support conceptual understanding (Chin & Brown, 2000; Wiebe Berry & Kim, 2008). Dialogic instruction necessarily needs more student voice, and student explanations are hypothesized here to predict higher reading outcomes.

Teacher and student explanations were each totaled and each divided by the number of minutes and words of each observation, and then averaged across the three observations per teacher to account for the differing lengths of transcripts. Explanation rates per minute that were classroom mean centered were used in the analyses.

First Language (L1) Use. Since bilingual students are over-sampled in this study, this code (L1Use) was included to capture any use of a Spanish-speaking student's first language to support student understanding or participation in instructional discourse. While not necessarily dialogic like an uptake question is, the use of a student's first language indicates an understanding of individual student linguistic resources which offers a scaffold to allow for more student voice in the lesson, particularly in linguistically-diverse classroom settings such as in the present study. Though current language policy in the Massachusetts district discourages the use of students' L1, it is expected that some use of Spanish by teachers proficient in the language will be observed due to the high proportion of Spanish-speaking students in both districts.

Instances of first-language use were coded and tallied per transcript, and rates per minute and per word were calculated and averaged across all observations of each teacher. The rates per minute were used in the analyses, and like the other codes, were classroom mean centered for the final models.

The full coding scheme, with the nine utterance level codes, four global level codes, and exemplars of each code are included in Appendix A.

Coding reliability.

Reliability was established at two levels of coding agreement: reliability on identifying all utterances that were codeable utterances, and reliability on each utterance-level code on the study's coding scheme (Appendix A).

First, agreement between two coders was established on the identification of codeable utterances in the transcripts, deciphering which instructional talk was managerial and what talk was instructional in content. Procedural, managerial and rhetorical utterances were not marked to be coded so that the analyses focused on the instructional moves that focused on and forwarded the instruction of literacy/language arts content, rather than on instructional moves that managed how the activity was enacted. For example, if a teacher asked "I'm sorry Amy, what did you just say?" this was not marked as a codeable question (say, Uptake) since this sort of brief clarification did not further the conversation or the lesson, but acted as a procedural move to help correct an utterance that was inaudible. Likewise, teacher and student read alouds directly from texts, common in these transcripts, were not marked for coding at the utterance level since a read aloud is not an instructional utterance with dialogic characteristics.

The reliability for marking what was codeable talk in each transcript was established by two research assistants to the larger CLAVES project, one at each research site. Both markers

were experienced classroom teachers and had observed roughly one-fifth of each site's lessons. To establish marking reliability between the two coders, seven randomly-selected transcripts (8% of total number of transcripts) were marked. The overall percentage of agreement was calculated by dividing the total amount of 'agreed-upon' markings by the total amount of agreed and discrepant markings. Reliability was established at an acceptably high 79%, with 70% or higher being a commonly accepted "rule of thumb" for many reliability coefficients (Fraenkel & Wallen, 2006, p. 161).

Once the transcripts were marked for what utterances to code, three coders coded 12.5% (transcript $n = 11$) of the total number of transcripts to establish reliability on the application of the coding scheme (Appendix A). Coder 1 had an extensive background in teaching elementary and middle school students in Massachusetts public schools. Coder 2 was a pre-service teacher in California preparing to teach secondary students in Social Studies, and a Spanish-English bilingual speaker. Coder 3 was doctoral student in Education with an extensive background teaching K-8 literacy in Maine public schools.

The three coders began by coding one transcript together, following the coding scheme closely. After agreeing on the coding, we coded another two transcripts separately. Coder 1 then compared each of these transcripts across the three coders, talking with each coder about their coding decisions. Each coder shared suggestions to refine and clarify definitions of the codes, provided more exemplars, and questioned how the codes were conceptualized in the context of the transcripts. Multiple iterations of the coding scheme were created and refined during this process. Using the finalized coding scheme, the three coders coded the 11 transcripts separately, and reliability was then calculated on these transcripts. Reliability was not established on the three global codes used to identify instructional materials, talk genres, or the field, since these

varied widely across all 88 transcripts and were coded using an inductive and emergent coding process by the first author.

To establish reliability among coders, an unweighted Cohen's Kappa (*K*) was calculated for each of the 11 utterance-level codes (four question types, two question quality codes, two types of explanations, two types of evaluations, and L1 use) and the one global code (teacher or student management of activities). Cohen's Kappa is appropriate for agreement on nominal scales and ordinal scales, indicating the proportion of agreement beyond that expected by chance (Sim & Wright, 2005). The default for codes in this coding scheme was zero, meaning that if a codeable utterance was not assigned a code, the feature was not observed by the coder. Thus, all codes were treated as binary, nominal codes to establish interrater agreement. The Kappa statistic indicates the "proportion of agreement beyond that expected by chance, that is, the *achieved* beyond-chance agreement as a proportion of the *possible* beyond-chance agreement" (italics in original; Sim & Wright, 2005, p. 258).

The Cohen's kappa coefficient is influenced by the prevalence of the code and the bias of rater disagreement on a given proportion of codes. A prevalence effect exists when the proportion of agreements on the existence of the code differs from the agreements on the non-existence of a certain code. When the prevalence index is large, the kappa is lower (Sim & Wright, 2005). Bias is the extent to which raters disagree on the existence or non-existence of a code, and affects the interpretation of the magnitude of the coefficient (Sim & Wright, 2005). When there is a large bias, the kappa coefficient is high, while the effect of bias is greater when the coefficient is small. Sim and Wright (2005), in their review of the kappa coefficient, suggest that the prevalence and bias indices be reported along with the coefficients when judging their magnitudes. In addition, the kappa coefficient does not account for sampling error; thus,

constructing a confidence interval can "indicate a range of plausible values for the "true" value of kappa" (Sim & Wright, 2005, p. 265). Generally, the standards for a kappa's strength of agreement are: ≤ 0 - .40 = poor – fair, .41 - .60 = moderate, .61 - .80 = substantial agreement, and .81 – 1 = almost perfect (Landis & Koch, 1977a; Sim & Wright, 2005).

Each instructional episode was coded for either teacher- or student-managed instruction (TM or SM, respectively). Field notes and teacher talk indicated the type of management emphasized in each instructional episode quite clearly, as indicated by a high kappa coefficient (κ = .88).

The inter-rater agreement reliability of utterance-level codes ranged from an unacceptably low kappa of .42 to perfect agreement. Authentic questions (AQ) resulted in the lowest and most unreliable kappa coefficient at .42, and quasi-authentic questions (QQ) were also only moderately reliable at κ = .57. AQ and QQ were dropped from the final models due to their low kappa statistics. Test questions indicated "substantive" rates of agreement between the three coders (κ = .61). Uptake questions showed the highest inter-rater agreement (κ = .68). Because there was very infrequent L1 use across all observed lessons, Spanish stood out and therefore scored perfect reliability. Reliability results are reported in Table 3.5, along with their respective prevalence and bias indices and confidence intervals.

Table 3.6 *Reliability (Cohen's κ) of Codes of Global and Utterance Features*

Code	κ	Prevalence index	Bias index	Confidence Interval
Global-level Codes				
Lesson Mgmt				
TM/SM	.88	.30	.01	.78 - .98
Utterance-level Codes				
Teacher Questions				
AQ^	.42	.90	.02	.33 - .50
QQ^	.57	.68	.004	.52 - .61

TQ	.61	.73	.004	.57 - .65
UP	.68	.75	.01	.64 - .72
Question Quality				
QuHigh	.89	.72	.03	.50 - .60
QuLow	.73	.34	.03	.70 - .76
Follow-up Moves				
TEvalH	.67	.75	.005	.63 - .71
TEvalL	.85	.48	.002	.82 - .87
TExp	.69	.89	.004	.62 - .75
Other Talk Moves				
StExp	.70	.99	.003	.52 - .89
L1Use	1.00	1.00	0.000	1.00 -1.00

Note. TM=teacher managed instructional episode; SM=student managed episode; AQ=authentic question; QQ=quasi-authentic question; TQ=test/display question; UP=uptake question; QuHigh=high quality question; QuLow=low quality question; TEvalH=teacher high quality evaluation; TEvalL=teacher low quality evaluation; TExp=teacher explanation; StExp=student explanation; L1Use=1st language use by teacher or student(s).
^ not included in final models

Second stage of qualitative analyses: Contextualizing the quantitative analyses.

The coding of transcripts was the first qualitative analysis of the data that then informed the quantitative stage of analysis (described below). These codes were converted to rates per minute and totaled, and then used as predictors in the multilevel models. In the second stage of qualitative analyses, the statistically significant findings from the models were explored to contextualize those talk moves that predicted students' reading comprehension outcomes (presented in Chapter 5). This recursiveness in the analytic process is a typical convergent mixed analytic design.

While coding, I concurrently kept track of my thinking about these data through memoing. I wrote memos to monitor the coding process and record any impressions about instructional moves and the instructional context. By doing this, I established an audit trail to establish some validity and reliability in this first stage of the qualitative analysis (Birks, Chapman, & Francis, 2008). This memoing led to the writing of "mini" case studies of each

classroom, documenting the learning activities set up by each teacher, the tone of the classroom as could be gleaned from the teacher's and students' comments, the teacher's behavioural management style, and other observations drawn directly from the data. Each observation included in these mini case studies directly referenced the data, staying as close as possible to the data. After coding all transcripts from each teacher, I reread the case studies of each lesson and wrote a one-page summary of each classroom, including my impressions on the general amount of dialogic instruction observed from the transcripts. The second level of qualitative analysis was driven by the quantitative findings but drew on these case studies and further readings of transcripts.

In order to harness the patterns and discrepancies of these instructional moves across 31 classrooms and 88 transcripts, each of the statistically significant moves were treated as cases in and of themselves. Following the work on case study analyses of Merriam (Merriam, 1998) and Stake (Stake, 2006), the talk moves of interest were treated as singular foci of analyses to explore their functions within classroom instruction and discourse that shaped the learning activities and demands experienced by students in these 31 classrooms. It is important to note that to my knowledge, individual talk moves have not been studied as cases in the extant literature, so while this analytic tact is unusual, it does provide a particularist picture of each case (Merriam, 1998; Stake, 2006), analyzing each talk move as a case in its "real-life context" of classroom instruction (Yin, 2009, p. 18). I consider each talk move a case because it is a "phenomenon … occurring in a bounded context" (Miles & Huberman, 1994, p. 25), each talk move being bounded by its lesson, and by its particularity compared to other moves during instructional talk. As Stake (2006) argues, a case is "a noun, a thing, an entity" rather than a verb or a function, and we understand a case by "experiencing the activity of the case as it occurs in

its contexts and in its particular situation," to find out "how the case gets things done" (p. 1-2). By presenting uptake questions, for instance, as a singular case, the quality of this question type can be explored to better understand how these questions might be precipitated by certain talk moves, or what patterns of talk might follow these questions. By looking more inductively at each of the five moves as they function within and across language arts lessons, the intent of this analysis is to consider how they function within patterns of talk to, at least partially, support students' reading comprehension achievement. Therefore, patterns and themes (Coffey & Atkinson, 1996) are identified in each case, and are analyzed across the sample of the 31 classrooms, as well as comparatively analyzed within classrooms characterized by high- and low- rates of the talk move. These qualitative analyses were managed using Atlas ti (Scientific Software Development, 2011), a qualitative software program. Coding frequencies, code proximities, and the co-occurrence of codes were explored to identify if the significant talk moves exhibited patterns across the transcripts, and what global or instructional features of lessons mediated these significant predictors. This process was fully documented by constant memoing on each query made with the software. This second stage of the qualitative analyses were undertaken to answer RQs 2 and 2a regarding the patterns and variations of the findings within and across classes.

Quantitative analytic approach.

After the transcripts were coded and reliability of coding agreement was established for each code, each code was counted and converted to rates per minute of instruction to account for varying lengths of classroom observations, as described previously. These variables were then merged with the student-level dataset to create a two-level dataset appropriate for multilevel analyses. The student-level dataset (or Level-1 dataset), comprised of the 236 students, included

reading and vocabulary assessment scores, and demographic variables. The teacher-level dataset (or Level-2 dataset), comprised of the 31 participating teachers, included all measures of instructional talk that met the acceptable levels of reliability in coding agreement, discussed previously.

Centering of Level-1 and Level-2 variables.

For ease in interpretation of the relationships between variables in the final hierarchical models, the student-level (Level-1) language and literacy measures and the classroom-level (Level-2) instructional talk measures were centered. Measures were centered on their respective grand means (i.e. the mean across all students and classrooms). Each measure's population average score was subtracted from each individual student's score on that measure to create a center of zero. When these variables are centered around the population means, β_{0j} (the intercept in the Level-1 model) represents the mean achievement level for the average student in classroom *j* (Raudenbush & Bryk, 1986). Centering creates a "meaningful 0-point" for an independent variable to allow for more ease of interpretation to compare an individual student's score to the average achievement score in the average classroom (Luke, 2004, p. 49). The Level-2 predictors (features of classroom instruction, in rates per minute) were centered on the mean across the 31 classrooms. Since each student in their respective classroom shares the same value on each of the Level-2 predictors, but the number of students per classroom varied from 2 to 12, I chose to center these variables by giving equal weight to each classroom. By centering on the mean across classrooms, effects can be interpreted as how much more or less of a given talk move an individual student is exposed to compared to the average for all students (Luke, 2004). Interpretation of the effects is based on classrooms being above or below average (represented by 0) on a given predictor. This calculation of centered, Level-2 variables is appropriate for

answering RQ1a where my substantive interest is in the effects of Level-2 predictors, controlling for individual literacy scores (Enders & Tofighi, 2007; Raudenbush & Bryk, 2002). The only two covariates in the final models that were not centered were the binary demographic variables for students' SES (farms) and first language status (firstling).

Model building.

Since these talk moves were observed at the classroom level, and students' reading comprehension abilities were measured at the student level, hierarchical linear modeling (HLM) was used to account for students clustered in classrooms and thus exposed to the same instructional talk moves with each teachers. Multilevel models were used to account for the correlation of errors in such clustered data. In addition, such models were appropriate because the proportion of between-classroom (level 2) variance in the dependent variable compared to the overall (between-student and between-classroom levels) variance was sufficiently high to warrant an HLM analysis (Luke, 2004), for a high intraclass correlation provides less unique information from each individual student (UCLA Statistical Consulting Group). An intraclass correlation coefficient showed that 20% of the variance of the WMLS-R Passage Comprehension standard score Wave 4 (PCSS4) test could be attributed to variance between classrooms. Stata (StataCorp, 2011) calculates the intraclass correlation variance by calculating the coefficient as the mean of the outcome variable (PCSS4) across the sample. The classroom-level variance term ($u = 23.70$) and the student-level variance term ($e = 91.11$) were combined ("fraction of variance due to u_i") to produce the intraclass correlation coefficient ($\rho = 0.20$).

The final HLM model was built using a step-wise approach. Only significant main effects were kept in the final model. Since the student sample size was acceptably high (student n = 236), variables were kept in subsequent models if they resulted in a p-value of 0.05; variables

below $p = 0.05$ were dropped from the models with the exception of the first language variable, since this is a variable of interest in the research questions guiding the study. First, a level 1 model was specified for the WMLS Passage Comprehension standard score at Wave 4 (PCSS4), *i*, for student *j*. This null, unconditional model was a multilevel model with no level 1 or 2 predictors:

Level 1: $Y_{ij} = \beta_{0j} + r_{ij}$

Level 2: $\beta_{0j} = \gamma_{00} + u_{0j}$

To control for students' initial reading and vocabulary levels, the Spring (Wave 3) reading and vocabulary measures were included as covariates to construct the best-fitting, Level 1 model as a baseline before constructing the classroom level (level 2) model. The WMLS Passage Comprehension standard score Wave 3 (PCSS3_PMC) population mean centered variable was the first covariate to be added to control for students' initial reading comprehension scores at the beginning of the school year. The other two WMLS tests, the Letter Word (LW3_PMC) and Picture Vocabulary (PV3_PMC) scores at Wave 3 were then entered together to this Level-1 model as population mean centered standard scores.

Next, the other language tests were entered together: the CELF Formulated Sentences test (FS3_PMC), the CELF Word Classes 2 test (WC3_PMC), and the morphological measure, Extract the Base (EB3_PMC), all as population mean centered scores, indicated by the _PMC annotation.

Students' demographic information was then entered as covariates to control for first language status (firstlang) and socioeconomic status (free and reduced price meals status; farms). As stated above, non-significant variables were not included in subsequent models; however, because bilingual students are of interest in this study, first language status was left in the models

despite non-significance at a .05 alpha level. These demographic variables were categorical variables: firstling (English = 0 and Spanish = 1), farms (non-farms status = 0, and farms = 1). Because standardized, age equivalent scores were used for the outcome variable (PCSS4), grade was not included as a covariate. The baseline, Level-1 model was thus composed of students' initial reading comprehension and vocabulary scores and their two demographic variables.

Once this appropriate baseline model was established, the nine Level-2 classroom instruction predictors were then entered. These covariates were rates per minute and classroom mean centered. As with the Level-1 covariates, classroom mean centered scores were calculated by taking the average of each of the nine classroom-level predictors across the 31 classrooms and subtracting these respective means from each student's nine classroom-level predictor scores. By using a total classroom centered mean, rather than a population (student *n*) centered mean, each classroom was given equal weight no matter the number of students in each classroom. As with the other covariates, these predictors were annotated with _CMC.

While it may make conceptual sense to group predictors by questions and follow-up moves, there is no theoretical reason for the order of entry of the predictor variables (ie. why enter questions in first, why not last?). Therefore, the nine Level-2 between-classroom predictors were entered together to create the following hierarchical linear model:

$$\text{Level 1: } PCSS4_{ij} = \beta_{0j} + \beta_{1j}(PCSS3\_PMC)_{ij} + \beta_{2j}(LWSS3\_PMC)_{ij} + \beta_{3j}(PVSS3\_PMC)_{ij} + \beta_{4j}(FS3\_PMC)_{ij} + \beta_{5j}(WC3\_PMC)_{ij} + \beta_{6j}(EB3\_PMC)_{ij} + \beta_{7j}(farms)_{ij} + \beta_{8j}(firstlang)_{ij} + r_{ij}$$

Level 2: $\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{TQ\_CMC})_j + \gamma_{02}(\text{UP\_CMC})_j + \gamma_{03}(\text{QUHigh\_CMC})_j + \gamma_{04}(\text{QULow\_CMC})_j + \gamma_{05}(\text{TEvalH\_CMC})_j + \gamma_{06}(\text{TEvalL\_CMC})_j + \gamma_{07}(\text{TExp\_CMC})_j + \gamma_{08}(\text{StExp\_CMC})_j + \gamma_{09}(\text{L12\_CMC})_j + u_{0j}$

The Level-1 model (above) specifies the relationships between the student-level variables: vocabulary and reading achievement, and demographics. In other words, this submodel distinguishes "the trajectories of different students using just their individual [achievement] parameters" (Singer & Willett, 2003, p. 54). The Level-2 model specifies the hypothesized population coefficients that demonstrate the effects of classroom-level variables on student achievement. Table 3.7 summarizes the parameters in the model.

Table 3.7 *Interpretation of Coefficients in the Final Model*

Variable		Variable Description
Level-1 Model		
Individual achievement parameters	β_{0j}	Intercept of mean achievement level for classroom *j*
	$(\text{PCSS3\_CMC})_{ij}$	Values of the student-level vocabulary, reading and demographic scores for student *i* in classroom *j*
	B_{kj}	Slope of student's, *i*, achievement based on the talk within classroom *j* (Regression coefficients that "characterize the structural relationships" within classroom *j* (Raudenbush & Bryk, 1986, p. 3))
Variance component	r_{ij}	Random error for student *i* in classroom *j*; Level-1 residuals
Level-2 Model		
Fixed effects	γ_{00}	Population average of Level-1 intercepts, β_{0j}; ie. the average achievement in the population; the mean intercept; the average adjusted mean
	γ_{0k}	Level-2 slopes; the average hypothesized difference in achievement between classrooms
	$(\text{TQ\_CMC})_j$	Predictors for classroom *j*

Variance component	u_{0j}	Random error at the classroom level; Level-2 residuals; "portions of intercepts and slopes leftover after accounting for the effects of … predictors" (Singer & Willett, 2003, p. 62)

The following two chapters are organized by the two main strands of the mixed analyses. Chapter 4 reports the quantitative results from the modeling described above, and Chapter 5 presents further qualitative analyses of the significant findings from these models.

Chapter 4 – Results: Quantitative Strand

This chapter first presents the descriptive statistics of covariates included in the HLM models to partially answer the first research question on the patterns and variations in dialogic instruction observed in the sample's classrooms. Each instructional code's quantity and range across the observation data provide a first look at the patterns and variations of the instructional characteristics in this upper elementary sample. The final HLM models and results are then presented to answer the research question of what features of instruction predict students' reading comprehension scores. Detailed qualitative analyses of the significant findings from the models are presented in Chapter 5.

The following research questions are addressed in this chapter:

RQ 1: What features of dialogic instruction predict reading comprehension scores?

RQ 1b: Are bilingual and monolingual students differentially impacted by exposure to high levels of dialogic instruction?

Preliminary Analyses: Patterns and Variations in Utterance Features of Instruction

The following tables and figures present descriptive statistics and correlations for the amount of teacher talk and questions, as well as the nine instructional talk codes applied to the classroom observation data. Nine codes were used to identify two kinds of teacher questions, two qualities of questions, two kinds of evaluations and explanations, and one code for L1 use during lessons. These nine features of instructional talk were used as classroom-level predictors in the multilevel models that follow this section.

Across the sample, the mean rate of total talk was 93.20 words spoken per minute (*SD*= 36.31), while teacher talk (total words minus student words in each transcript) averaged 77.86 words spoken per minute (*SD*= 30.33). Therefore, teacher talk represented 84% of the talk during

these lessons. Within this talk, teachers across the sample posed an average of one question per minute (1.07, *SD*= .45). Table 4.1 presents these overall talk rates disaggregated by grade level.

Table 4.1 *Mean (SD) Rates of Talk, Teacher Talk and Questioning by Grade*

	Overall Rate of Talk		Teacher Rate of Talk		Teacher Question Rate	
Grade	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
Grade 3	87.50 (38.52)	33	77.23 (32.81)	33	0.98 (0.47)	33
Grade 4	96.67 (32.64)	33	79.76 (26.97)	33	1.32 (0.66)*	33
Grade 5	96.56 (38.69)	22	75.96 (32.42)	22	0.85 (0.40)	22

Note. Rates of talk are presented as audible words per minute. Teacher question rates reflect audibly instructional questions per minute. N represents the number of transcripts used for analysis.
* Significant difference between grades 4 and 5.

As Table 4.1 indicates, average teacher talk rates were consistent across the grade levels, though with such high standard deviations, it is clear that teacher talk varied considerably between classes. Average teacher questioning rates were not as similar across grades, with fifth-grade teachers posing fewer questions, on average, than the other two grades. The average rate of teacher questions was significantly different between grades four and five ($t(53) = 2.90$, p=.0027). Two-tailed, two sample t-tests comparing the teachers' rate of questioning and overall rates of talk between grades 3 and 4, and grades 3 and 5 showed no significant difference between populations.

Tables 4.2 and 4.3 present raw count data and average rates of each code. Some codes, like teacher and student explanations (TExp and StExp, respectively), and first-language use (L1Use) were used infrequently, and these small numbers were characterized in particular by non-normal distributions. This issue is addressed below.

Table 4.2 *Mean Rates (SD) of Raw Counts of Instructional Codes across the Sample*

Variable Name	Mean (SD)	Min	Max
Test/display questions (TQ)	17.99 (14.85)	0	68
Uptake questions (UP)	17.55 (13.84)	0	80
High-level questions (QUHigh)	17.81 (13.02)	0	80
Low-level questions (QULow)	39.86 (23.47)	1	104
Teacher Evaluations, High (TEvalH)	19.42 (13.13)	0	64
Teacher Evaluations, Low (TEvalL)	30.53 (20.00)	1	104
Teacher Explanations (TExp)	7.23 (5.61)	0	33
Student Explanations (StExp)	9.58 (9.39)	0	34
First-language use (L1Use)	0.42 (1.83)	0	13

Note: N = 88 transcripts; these averages do not control for length of observations

The averages in Table 4.2 are based on total numbers of codes applied across the 88 transcripts, so they do not take into account the length of transcripts and thereby do not reflect the fact that in most transcripts, the longer the transcript (ie. the observation), the higher the total number of codes applied. Table 4.3 presents the average rates of each code per minute of transcript to account for the differing lengths of the classroom observations. These mean rates of each instructional code are also averaged for each teacher across the three transcripts collected in each classroom. Test, quasi-authentic, and uptake questions were the most frequently used across the sample, along with low-level teacher evaluations.

Table 4.3 *Classroom Mean Rates per Minute (SD) of Teacher Talk Moves across the Sample*

	Mean Rate (SD)	Skewness	Kurtosis	Min	Max
Teacher Questions					
Overall Rate of Questioning	1.05 (0.44)	0.82	3.63	0.29	2.22
Test Questions	0.33 (0.19)	0.97	4.15	0.06	1.00
Uptake Questions	0.32 (0.19)	1.78	6.59	0.09	1.00
High-quality Questions	0.33 (0.18)	2.16	9.14	0.05	1.01
Low-quality Questions	0.73 (0.31)	0.66	2.62	0.25	1.60
Follow-up Moves					
High-quality Evaluations	0.33 (0.17)	0.57	3.58	0.05	0.88
Low-quality Evaluations	0.54 (0.22)	0.86	4.29	0.16	1.35

Teacher Explanations	0.12 (0.08)	0.78	2.67	0.02	0.33
Student Explanations	0.16 (0.12)	0.56	2.41	0.00	0.44
L1 Use	0.01 (0.02)	2.75	9.24	0.00	0.08

Note. Mean rates = rate of talk move per minute of instruction; Sample = 236 students

The skewness and kurtosis statistics indicated that all the variables of talk moves had distributions with slight to strong positive skews; L1 Use was the least-occurring code and had the strongest positive skew. The value of kurtosis for a normal distribution is 3 (Stata software uses this calculation), and variables in Table 4.3 have kurtosis around 3, indicating fairly normal peaks to the distributions; the exceptions were the test questions (TQ; $p = .01$), uptake questions (UP; $p < .001$), high question quality (QUHigh; $p < .001$), low-quality evaluations (TEvalL; $p = .01$) and first language use (L1Use; $p < .001$) distributions. These five variables showed a statistically significant lack of normal skewness and kurtosis. While these significant findings are a concern for assumptions of normality, it is worth noting that a kurtosis with an absolute value of 10 or higher is considered problematic (Acock, 2010). Despite this, these descriptive statistics make it clear that these variables cannot be assumed to have normal distributions, which may increase the likelihood of committing a Type II error (ie. failing to reject the null hypothesis when null is false). Dichotomizing some variables is one approach to address this potential problem. This is considered in each of the following sub-sections describing the distributions of the nine Level-2 codes in more detail below.

Teacher questions.

The most common type of question used by teachers in all three grades was test questions (TQ), with an average rate of 0.33 ($SD = 0.19$) TQs per minute across the grades (Table 4.3). TQs were teacher questions that presumed students knew the answer because it was in the text or was addressed earlier during instruction, and had only one possible correct answer (Applebee et

al., 2003; Boyd & Rubin, 2006; Nystrand et al., 2003). Also frequent across the grade levels were uptake questions (UP; mean = 0.32, *SD* = 0.19), teacher questions that incorporated a student's previous contribution (e.g. *Why? Tell me more*). In addition, both types of questions were coded for what kinds of thinking they elicited. The cognitive quality of questions were consistently low-level (QULow), with a mean rate of 0.73 (*SD* = 0.31). The answers to these questions were found in the text or in previous utterances and explanations; in essence, these low-level questions sought answers drawn from the routine application of prior knowledge (e.g. *What happened to the main character? What is this? What does 'consider' mean? Or, fill-in-the-blank questions*).

There was variability in the quantity of teacher questioning between grades three, four, and five (Table 4.4). Fifth-grade teachers posed less questions per minute of transcript on average than the other two grades, and significantly less questions than grade four classes ($t(17) = 2.24$, $p < .05$). Two-tailed t-tests comparing the teachers' rate of questioning between grades 3 and 4, and between grades 3 and 5 showed no significant differences. In the fourth and fifth grades, uptake questions were posed more frequently than other types of questions, while third-grade teachers posed TQs most frequently. When the rates of question types were compared, a pattern of difference between the fourth and fifth grade samples emerged. Fourth-grade teachers used significantly higher rates of test questions ($t(17) = 1.73$, $p < .05$) and uptake questions ($t(17) = 1.81$, $p < .05$) than fifth-grade teachers. But fourth-grade teachers also used significantly more low-quality questions than in the fifth-grade ($t(17) = 2.63$, $p < .01$). There were no significant differences in question quality between the third and fourth or third and fifth grades.

Table 4.4 *Mean (SD) Rates of Questioning, Question Types, and Question Quality, by Grade*

	Grade 3	Grade 4	Grade 5

Measure	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
Overall Rate of Questioning	0.97 (0.38)	12	1.32 (0.53)	11	0.84 (0.30)	8
Teacher Question Types						
Test Questions	0.35 (0.20)	12	0.39 (0.25)	11	0.22 (0.14)	8
Uptake Questions	0.24 (0.12)	12	0.43 (0.24)	11	0.27 (0.10)	8
Question Quality						
High	0.28 (0.12)	12	0.39 (0.23)	11	0.31 (0.16)	8
Low	0.70 (0.30)	12	0.92 (0.36)	11	0.55 (0.21)	8

Note: Talk moves were measured as audible instructional questions per minute

A more general trend showed a preponderance of low-level quality questions (QULow) versus high-level questions (QUHigh) at each grade level, suggesting that students were asked to recite or report known information more frequently than asked to infer, generalize, predict or evaluate information. Similar to assertions in the literature that test/display/recitative questions tend to be lower-level and less dialogic (Lotman, 1988; Nystrand et al., 1997; Nystrand et al., 2003), low-level quality questions were highly associated with test questions ($r = .83$, $p < .001$). Uptake questions were shown to have strong correlations with both high-quality and low-quality questions ($r = .80$, $p < .001$ and $r = .75$, $p < .001$, respectively). However, these two constructs, categorizing questions as high or low quality, may not be as indicative of differences in the quality of instruction since the two constructs were also moderately correlated ($r = .53$, $p < .001$). Table 4.9 reports all correlations between level-1 and level-2 predictors.

The correlations between TQ, Up, and question quality are presented visually in Figures 4.1 and 4.2. With the highest correlation at $r = 0.83$, $p < .001$, it is clear that test questions have a stronger relationship with lower-quality cognitive activities than high quality questions ($r = 0.30$, $p < .001$). A scatterplot of the relationship and fit lines between test questions and high and low question quality demonstrate this stronger, positive relationship between low-quality questions and test questions (Figure 4.1). The mean rates of test questions across the 31 classrooms are represented on the x-axis, while the y-axis shows the mean rates of high and low question quality

codes. The second highest correlation between Level-2 variables was between uptake questions and the high-quality question code, but as can be seen in Figure 4.2, there is a less distinct relationship between uptake questions and high or low question quality across all grade levels, than with test questions. However, this strong relationship between Up questions and high-quality questions ($r = 0.80$, $p < .001$) is important in the final models, as they make indicate a suppression effect. This will be discussed in the discussion of the results in Chapter 6.

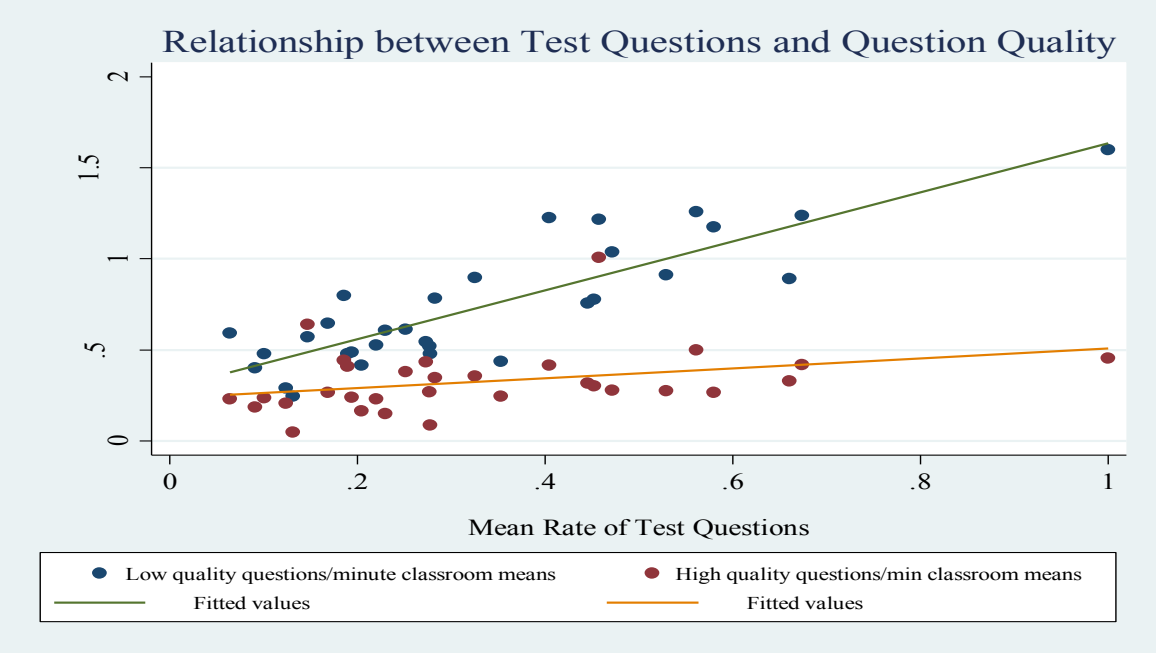


Figure 4.1 *Relationship between Test Questions (TQ) and Question Quality Rates*

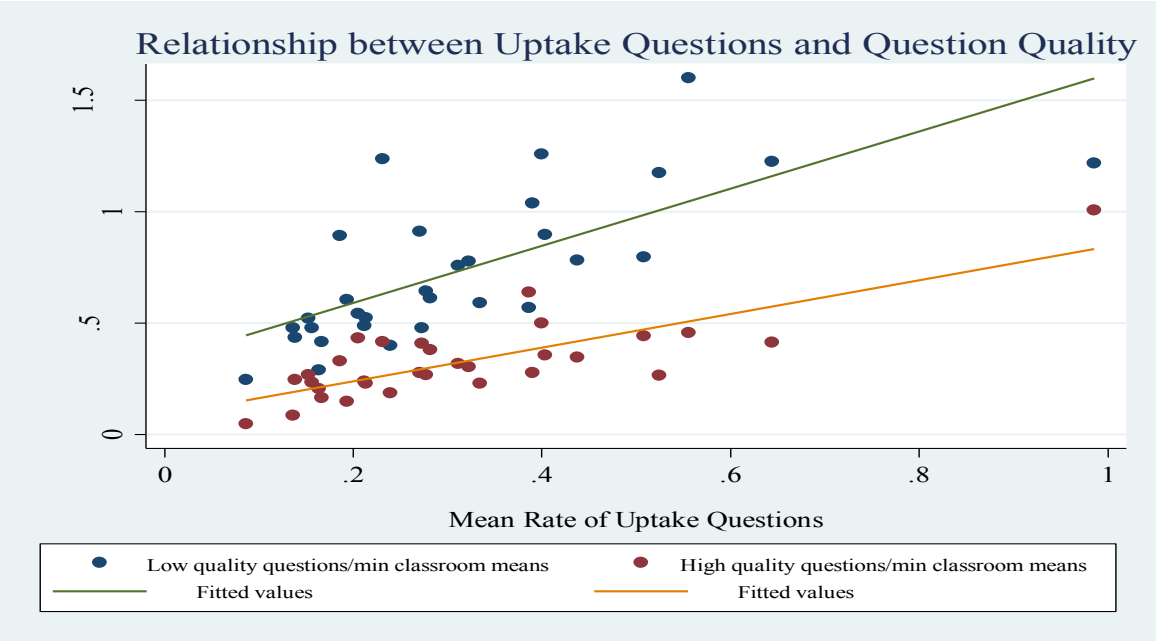


Figure 4.2 *Relationship between Uptake Questions (UP) and Question Quality Rates*

Introduced above, the distributions of each question type were positively skewed, but TQs, Ups, and High-level questions showed significant positive skews. In the following figures, the characteristics of the distributions of the mean rates of each type of question are graphed. The distribution of the average rates of test questions (TQs) used during instruction was fairly normally distributed with a positive skew (Figure 4.3). A scatter plot of student reading comprehension scores (WMLS-PC standardized scores) and exposure to higher rates of TQs shows a moderately positive relationship (Figure 4.4). As the TQ rate did not show a floor effect, and the distribution, though not ideal, was fairly normal, this predictor was kept as a continuous variable in all models.

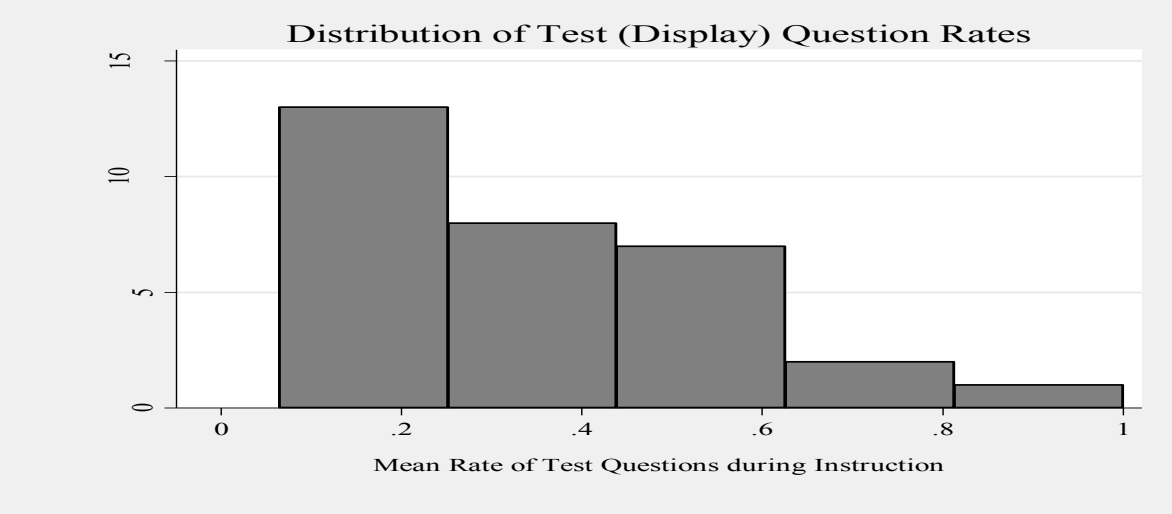


Figure 4.3 *Distribution of Test (TQ) Question Rates*

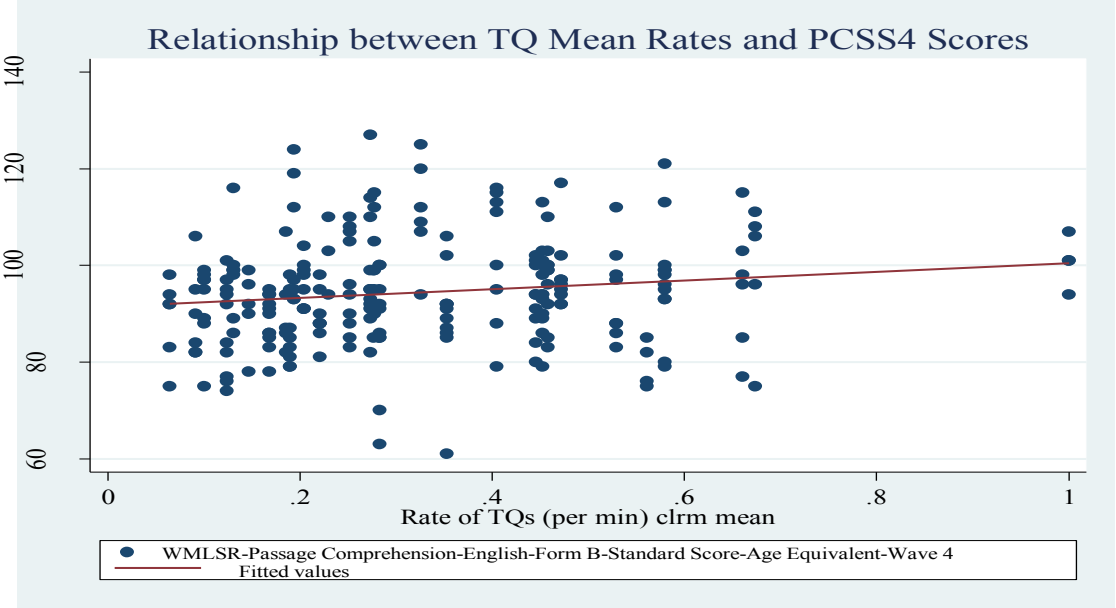


Figure 4.4 *Relationship between TQ Mean Rates and PCSS4 Scores*

Likewise, the distribution of the average rates of uptake questions (UP) across the 31 classrooms was not an ideal normal distribution, but significantly positively skewed ($p < .001$; Figure 4.5). The relationship between this type of question rate and student scores was the most

positively related, as indicated by the scatter plot in Figure 4.6. Though not ideally distributed, this UP rate variable was left as a continuous variable in the final models.

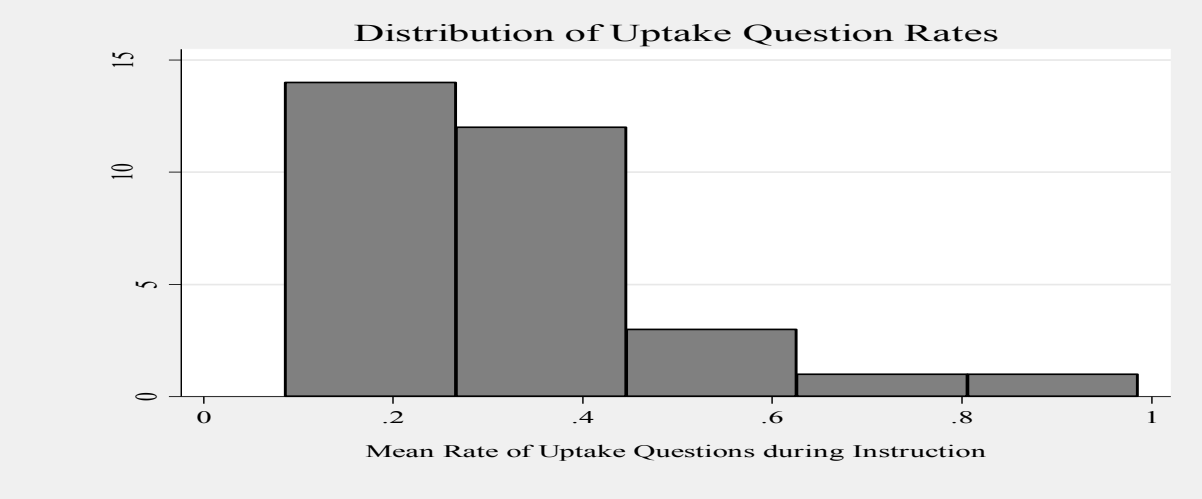


Figure 4.5 *Distribution of Uptake Question (UP) Rates*

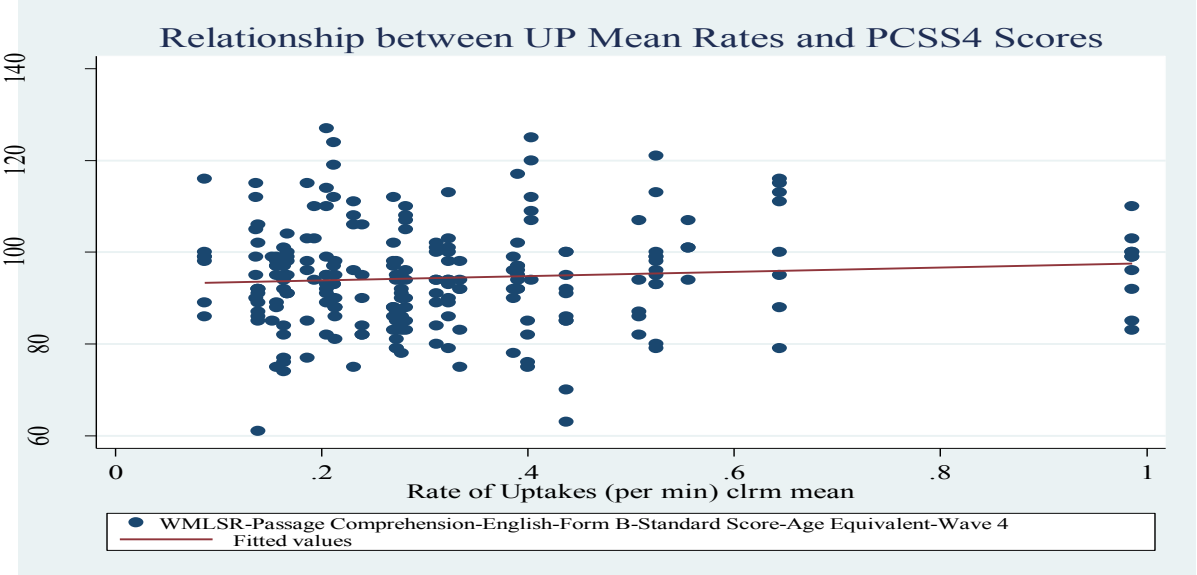


Figure 4.6 *Relationship between UP Mean Rates and PCSS4 Scores*

High- and low-quality questions were the most normally distributed variables of all the 'question' codes applied to the transcript data. Figures 4.7 and 4.8 show their respective distributions. Both predictors were continuous variables in the analyses.

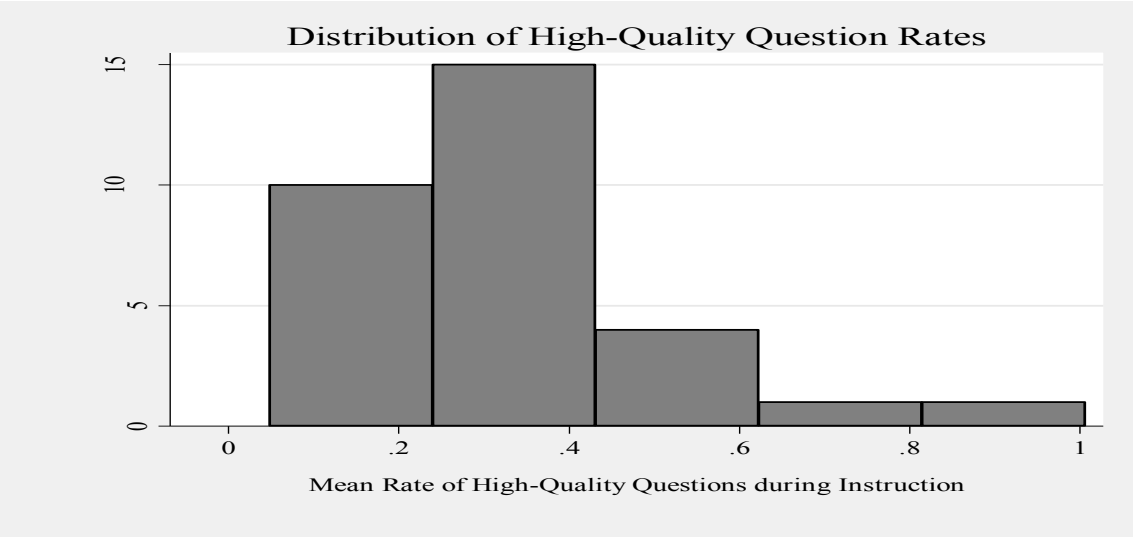


Figure 4.7 *Distribution of High-Quality Question Rates*

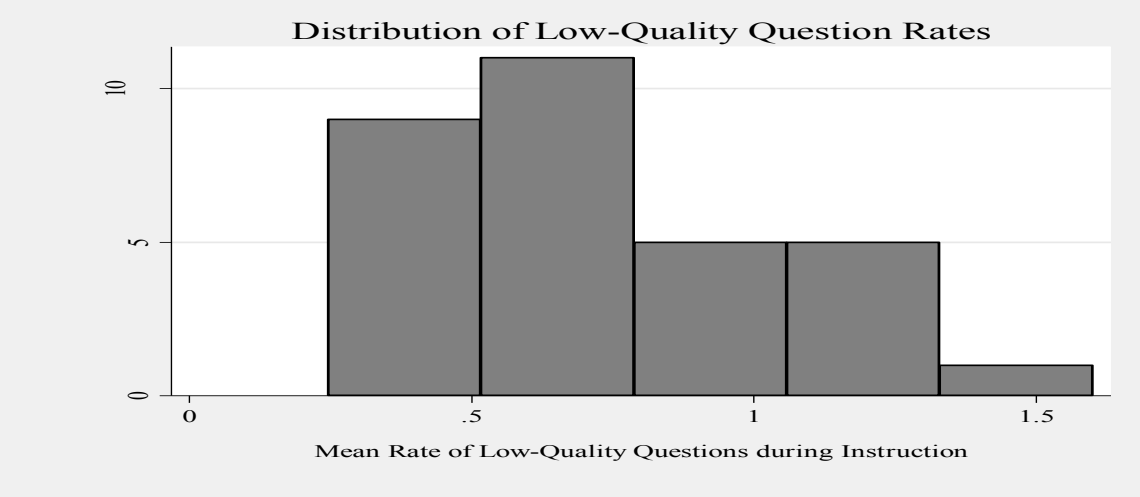


Figure 4.8 *Distribution of Low-Quality Question Rates*

Scatter plots of both kinds of question quality codes indicate that low-quality questions had a more positive relationship with student WMLS-PC standardized scores than high-quality questions, which would seem counterintuitive to current theories of reading instruction (Figures 4.9 and 4.10). Pearson correlations are presented in Table 4.9 below.

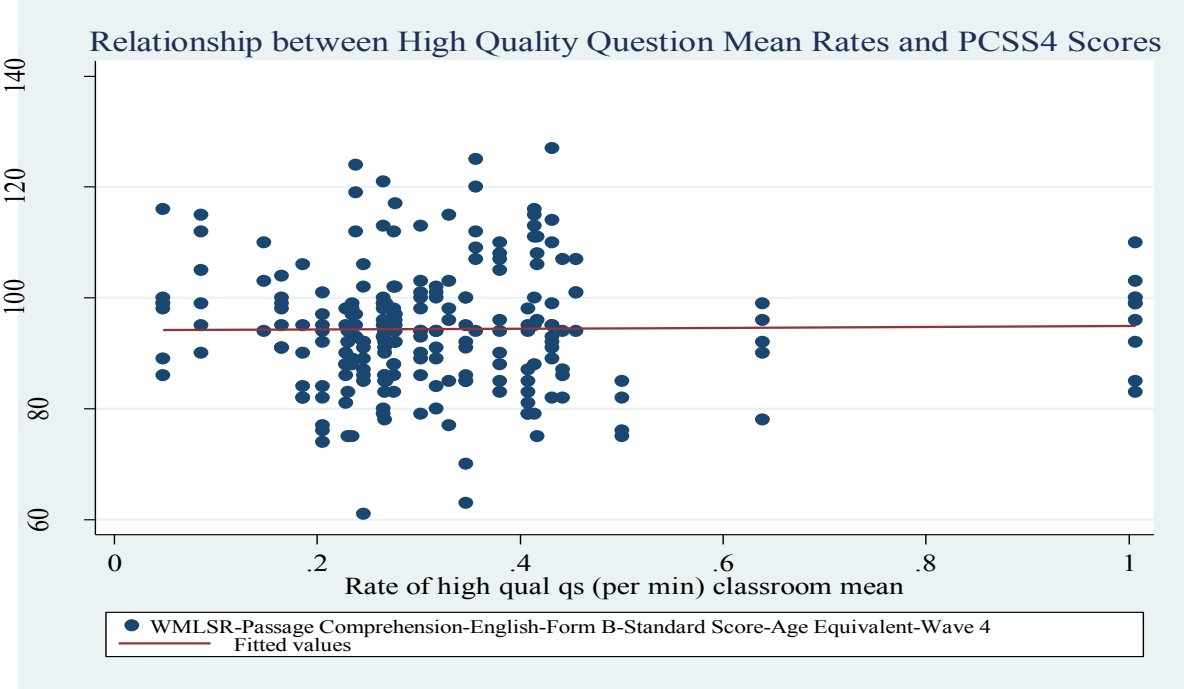


Figure 4.9 *Relationship between High-Quality Question Mean Rates and PCSS4 Scores*

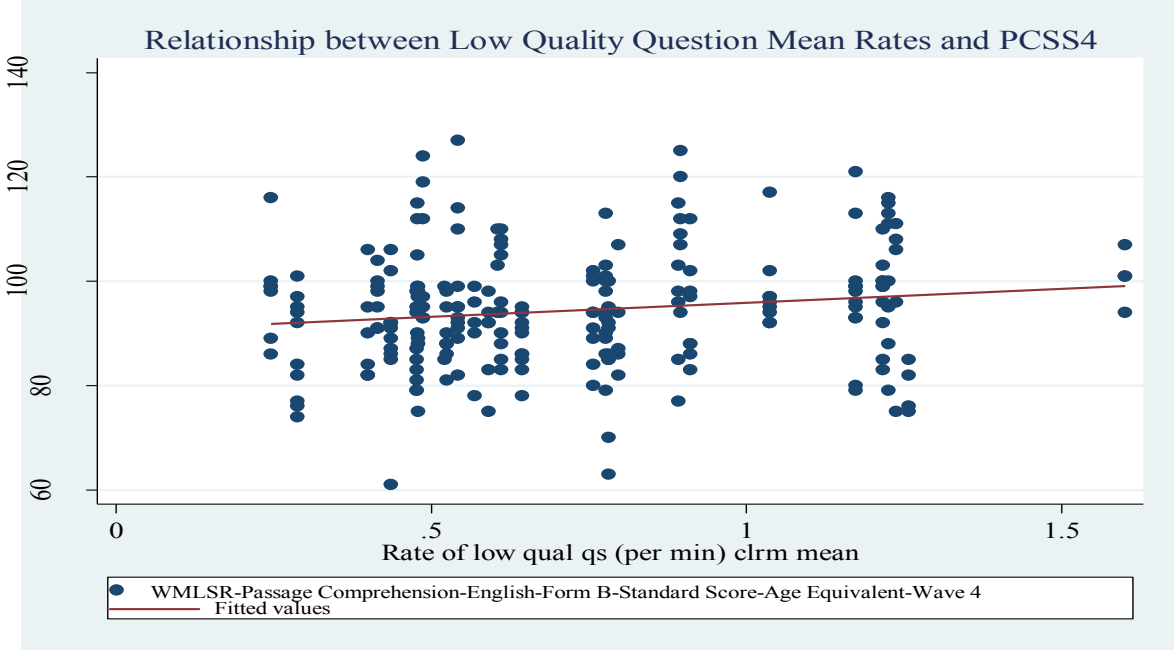


Figure 4.10 *Relationship between Low-Quality Question Mean Rates and PCSS4 Scores*

To summarize the patterns and variations of teacher questions across the 31 classrooms (RQ1), there were no significant differences of talk rate between the grades, but there was a difference between the rate of questioning and the kinds of questions posed between fourth- and fifth-grade classes. The fifth-grade teachers did not pose as many questions as their colleagues, nor as many TQ and Up questions as their fourth- and third-grade colleagues. Across all three grades, test questions and low-level quality questions were most frequently used on average. Finally, while most distributions were positively skewed, they proved to be within an acceptable normal range and were left as continuous variables in the final models.

Teacher evaluations and teacher/student explanations.

Teacher follow-up moves of student contributions were coded to assess their quality of dialogic instruction. Understanding the function of teacher questions drew on understanding what responses teachers were looking for by considering teacher evaluations to students'

responses. This "E" move in the IRE pattern has not been well studied, but some researchers have suggested that this move is potentially more important for student learning than the kinds of questions teachers pose (Nassaji & Wells, 2000). Moves that "revoice" (O'Connor & Michaels, 1993) or "recast" (Sharpe, 2008) student contributions were considered 'high-level' evaluations (TEvalH) because they focused the class's attention, for the moment, on the student's contribution and provided support for content or language learning. For instance, when a student tries to define how an author uses the word 'awesome', the teacher responds: "It can describe an emotion. It can describe how you feel about something. In looking at that mountain and looking at the beautiful fir trees on it, … the snow at the top, …. In that case I'm talking about its beauty, right. We can use that word [awesome] in many ways" (P83:35, 2/18/11). 'Low-level' evaluations (TEvalL) coded teacher utterances that either gave a brief recognition of the student's answer (e.g. "good, that's right," "uh-huh," or a verbatim repetition of the contribution), or gave no response (e.g. the teacher immediately calls on another student to contribute, or is distracted and does not ask the student to repeat his/her answer).

Along with teacher evaluations, teacher explanations (TExp) were counted to indicate the frequency of uninterrupted instructional talk teachers used. Explanations were utterances that forwarded the lesson by explaining a concept without student input and without posing questions over a minimum of two lines of transcript, or talk that included two or more ideas. Teacher explanations were thought of as 'stand alone' instructional talk, or proxies for direct instruction like lecturing. Conversely, the code 'student explanations' (StExp) was applied to all student utterances that contained a minimum of two lines of transcript and expressed a coherent idea. This code was used as an indicator of the amount and frequency of extended, on-task student talk

during lessons. Student writing that was shared aloud with a teacher, peers, or with the class, was also coded as StExp.

Overall, teachers more frequently responded with low-quality than high-quality evaluations across the grades. Table 4.5 presents the descriptive statistics of how teachers in this sample responded to students' answers and contributions. Teachers in grade four showed the highest mean rates of both types of evaluations, and was statistically significantly larger than the mean TEvalL rates used in the fifth-grade ($t(17) = 1.90$, $p < .05$).

Table 4.6 shows the average frequency of teacher and student explanations, sorted by grade level. There were very few teacher explanations per minute at all three grade levels, and all indicated large standard deviations. The average frequencies for student explanations were more variable between the grades, with third-grade students showing a menial average of 0.09 (*SD* = 0.06) of extended talk instances per minute of instruction, while fourth- and fifth-grade students were observed to use explanations more frequently than third-grade students, and more frequently than their respective teachers, on average, though again with large standard deviations. These rates between the three grades were not significantly different from each other. Overall, there were few StExp codes applied to the transcripts, which can partially account for the wide variations around the means.

Table 4.5 *Mean Rates of High- and Low-Quality Teacher Evaluations, by Grade*

	Teacher Evaluations							
	High-quality				Low-quality			
	Mean (SD)	N	Min	Max	Mean (SD)	N	Min	Max
Grade 3	0.26 (0.14)	12	0.05	0.55	0.51 (0.19)	12	0.28	0.96
Grade 4	0.45 (0.12)	11	0.14	0.64	0.69 (0.28)	11	0.33	1.35
Grade 5	0.33 (0.25)	8	0.10	0.88	0.47 (0.20)	8	0.16	0.76

Note: N = classrooms

Table 4.6 *Mean Rates of Teacher and Student Explanations, by Grade*

	Explanations							
	Teacher				Student			
	Mean (SD)	N	Min	Max	Mean (SD)	N	Min	Max
Grade 3	0.16 (0.11)	12	0.03	0.33	0.09 (0.06)	12	0.01	0.19
Grade 4	0.13 (0.05)	11	0.05	0.22	0.20 (0.14)	11	0.00	0.38
Grade 5	0.11 (0.09)	8	0.02	0.26	0.24 (0.13)	8	0.05	0.44

Note: N = classrooms

In the following figures, the characteristics of the distributions of the mean rates of each follow-up talk move are graphed. The frequency distributions of high- and low-quality teacher evaluations were fairly normally distributed with slightly positive skews (Figures 4.11 and 4.13, respectively). The kurtosis statistics showed reasonable heights to each distribution. The distribution of high-quality evaluations spoken per minute (Figure 4.11) did not show floor effects, so every teacher provided at least one high-quality evaluation of a student response during each observed lesson. Two-way scatter plots (Figures 4.12 and 4.14) of these average rates of high- and low-quality evaluations with students' wave 4 WMLS- PC standardized scores indicated positive relationships, with the plot between TEvalL and students' outcome reading comprehension scores showing a strong, positive relationship (Figure 4.14). For these reasons, TEvalH and TEvalL were left as continuous variables in the final models.

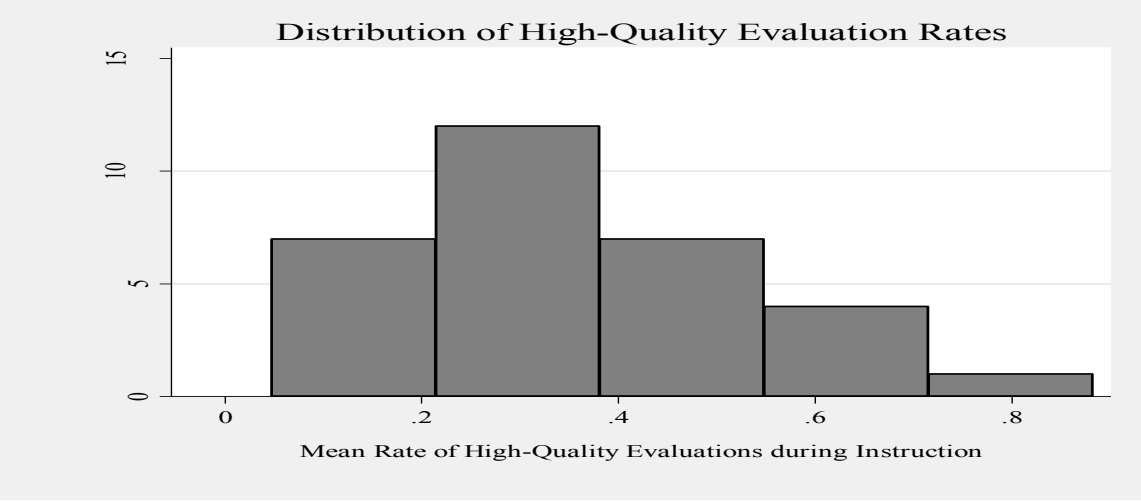


Figure 4.11 *Distribution of High-Quality Evaluation (TEvalH) Rates*

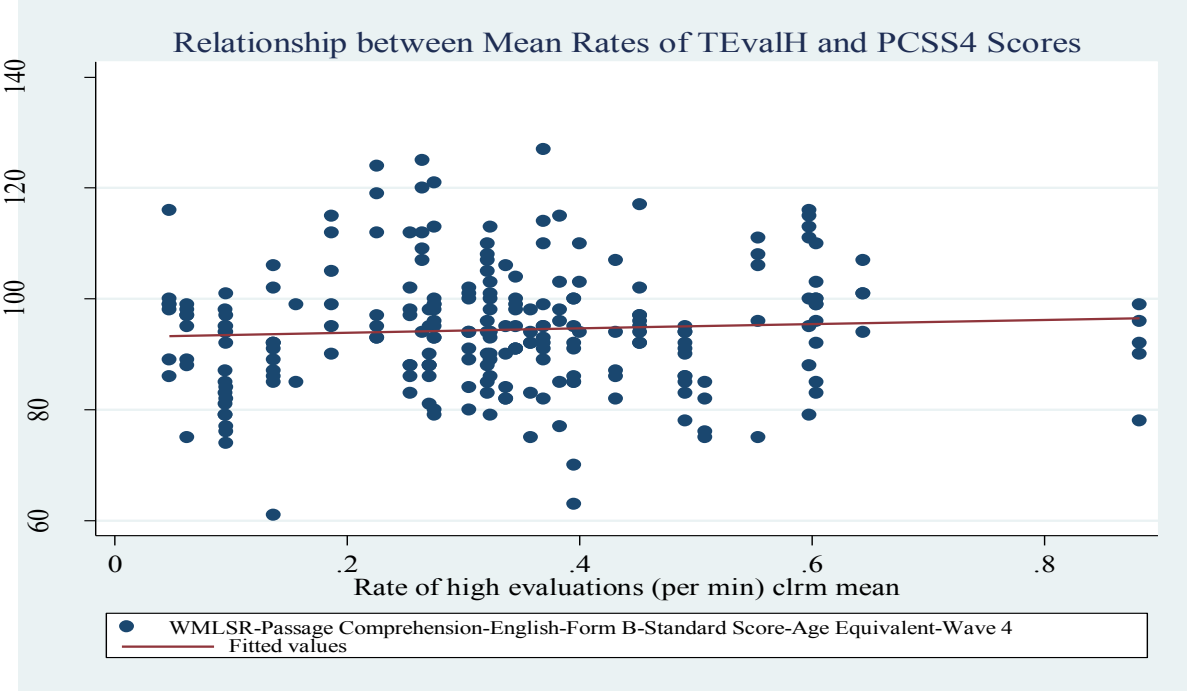


Figure 4.12 *Relationship between Mean Rates of TEvalH and PCSS4 Scores*

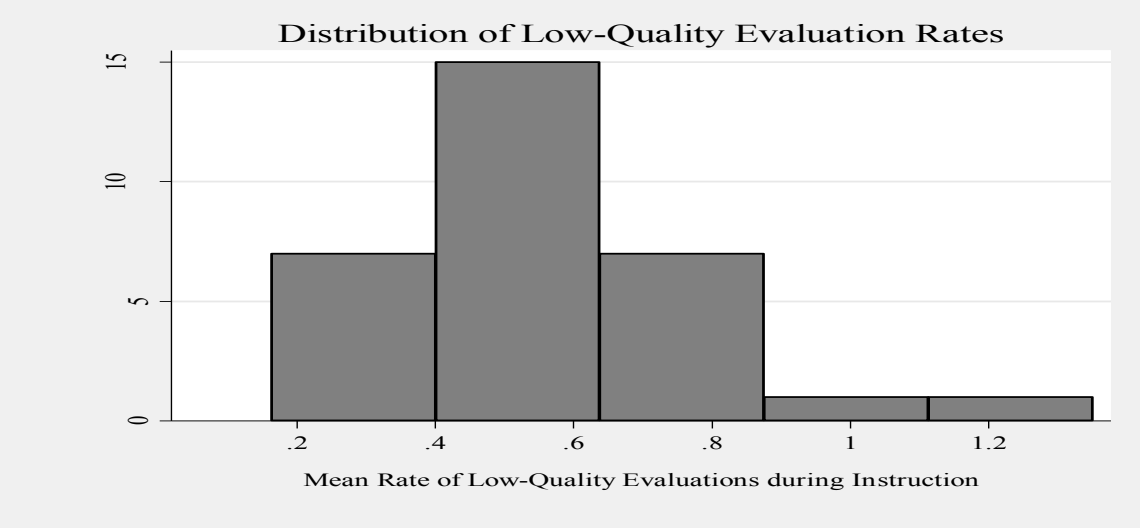


Figure 4.13 *Distribution of Low-Quality Evaluation (TEvalL) Rates*

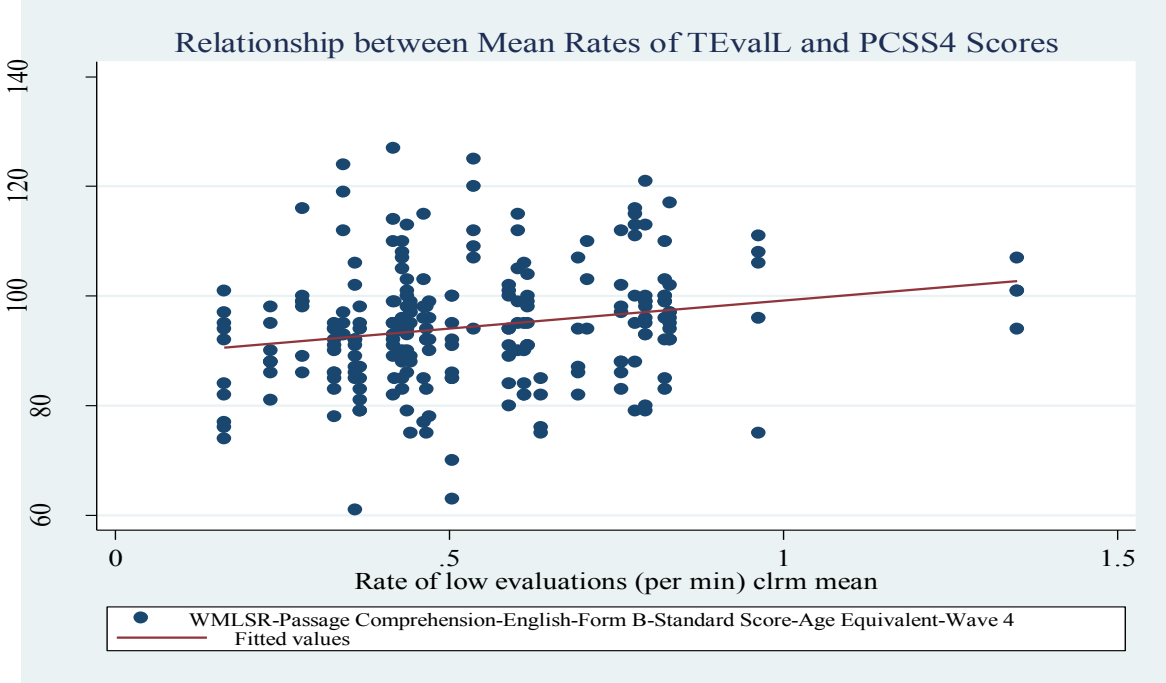


Figure 4.14 *Relationship between Mean Rates of TEvalL and PCSS4 Scores*

Rates of teacher and student explanations showed less normally distributed means. Teacher explanations were slightly positively skewed (Figure 4.15) and indicated a weakly positive relationship with students' PCSS4 standardized scores (Figure 4.16). Student explanations showed a similarly positively skewed distribution (Figure 4.17) and a flat relationship between student outcome scores (Figure 4.18). Pearson *r* values between these variables are presented in Table 4.9 below.

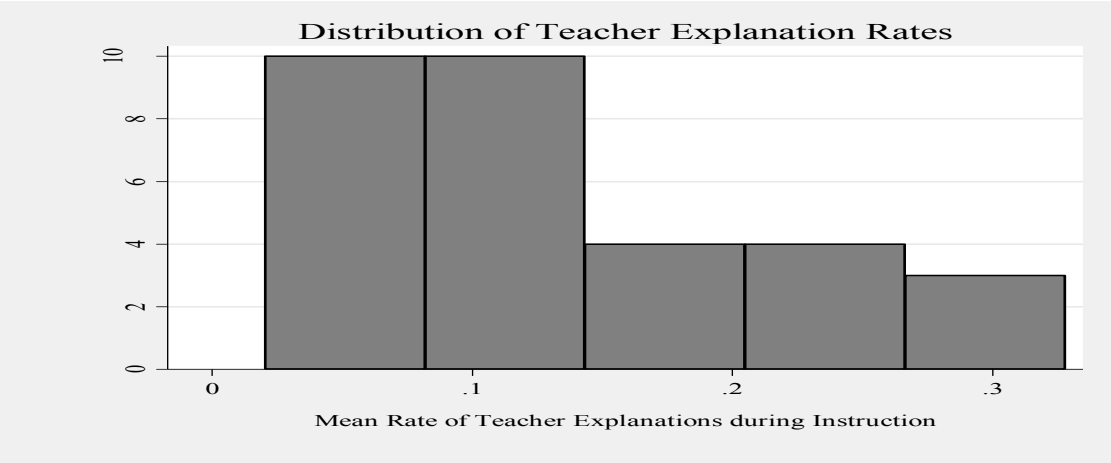


Figure 4.15 *Distribution of Teacher Explanation (TExp) Rates*

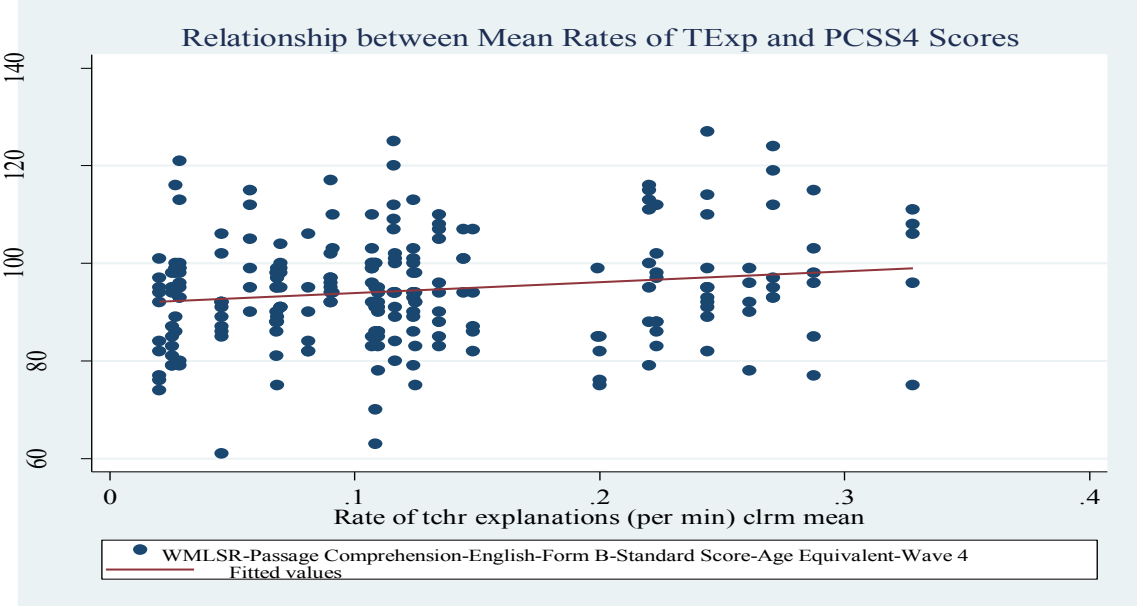


Figure 4.16 *Relationship between Mean Rates of TExp and PCSS4 Scores*

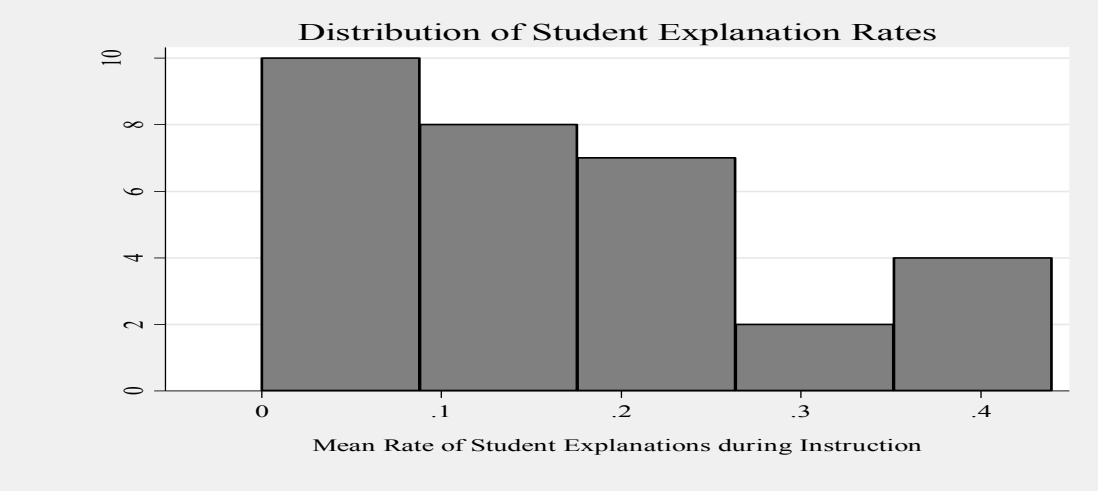


Figure 4. 17 *Distribution of Student Explanation (StExp) Rates*

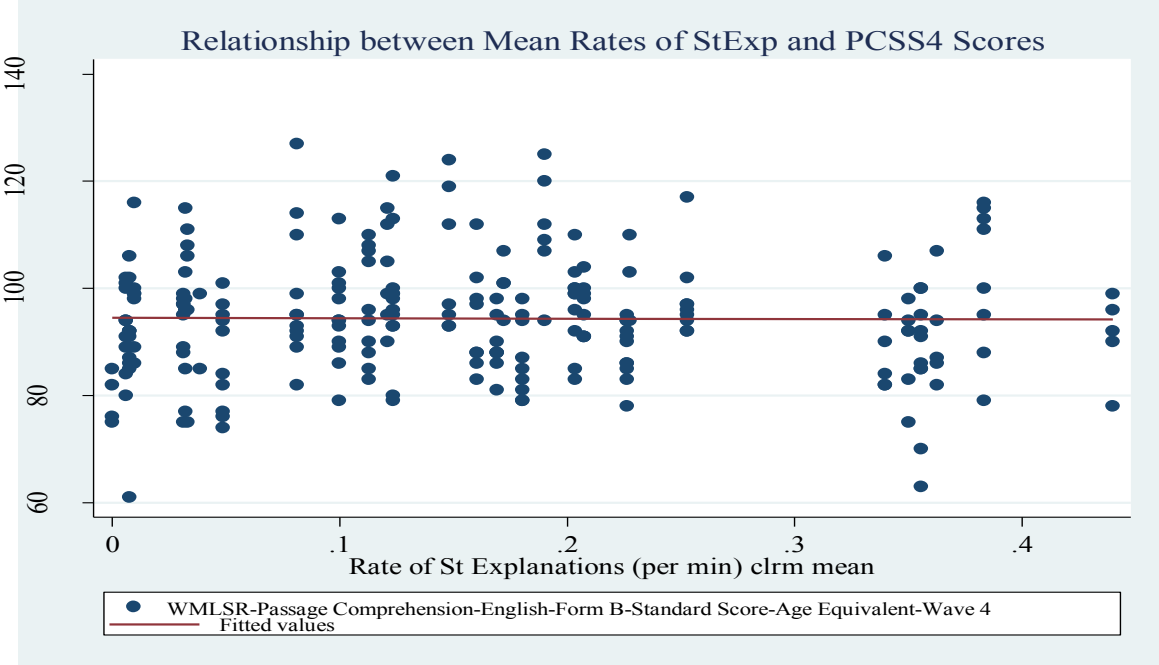


Figure 4.18 *Relationship between Mean Rates of StExp and PCSS4 Scores*

First-language use.

The final code applied to utterances in the observational data was first-language use (L1Use). This code captured each time a student's first language (Spanish in all cases here) was used to aid instruction. There were very few instances of Spanish language use, and most

instances occurred in the only Sheltered English Instruction (SEI) classroom in the sample (newly immigrated students in Massachusetts). Their fourth-grade teacher used Spanish to translate directions and some content during brief, one-to-one conversations with a few students. Table 4.7 shows the descriptive statistics for the classroom mean rates (per minute) of this L1Use code across the grades. As so few L1Use codes were applied to third-grade instruction, descriptives are reported to three decimal points. Fourth- and fifth-grade teachers used the highest rates of Spanish during instruction, though clearly, the range of this feature of instruction exceeded their means, suggesting a non-normal distribution to these data. There were no statistically significant differences of L1Use between the grades.

Table 4.7 *Average Rates (SD) of First-language Use, by Grade*

	Mean (SD)	N	Min	Max
Grade 3	0.002 (0.004)	12	0	0.01
Grade 4	0.01 (0.03)	11	0	0.07
Grade 5	0.01 (0.03)	8	0	0.08

Note: N = Total number of classrooms

The histogram in figure 4.19 clearly shows this code was not normally distributed. Most classrooms did not use any first language during instruction, showing a floor effect on this predictor. While transforming the variable would normalize the distribution, because of the preponderance of zeros, L1Use was dichotomized. Therefore, L1Use was dummy coded so that 0 = no L1Use, and 1 = some L1Use occurred during instruction. Figure 4.20 shows the relationship between the PCSS 4 with the continuous L1Use variable, showing a slightly negative relationship between the two. Figure 4.21 shows the same relationship between the outcome reading scores and the dichotomous L1Use variable. This dummy variable was used in the final models.

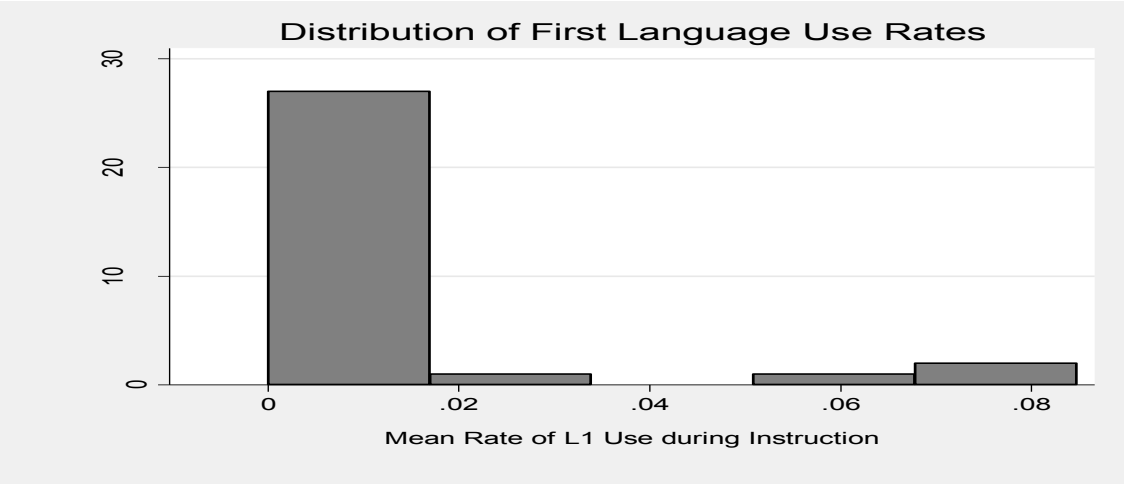


Figure 4.19 *Distribution of First Language Use (L1Use) Rates*

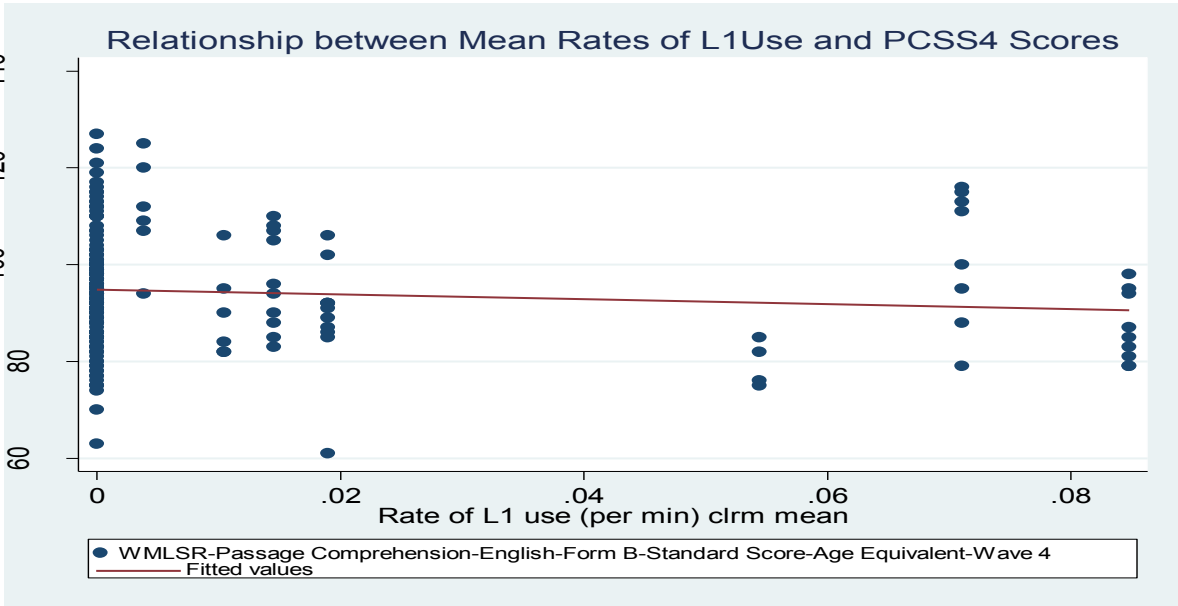


Figure 4.20 *Relationship between Mean Rates of L1Use and PCSS4 Scores*

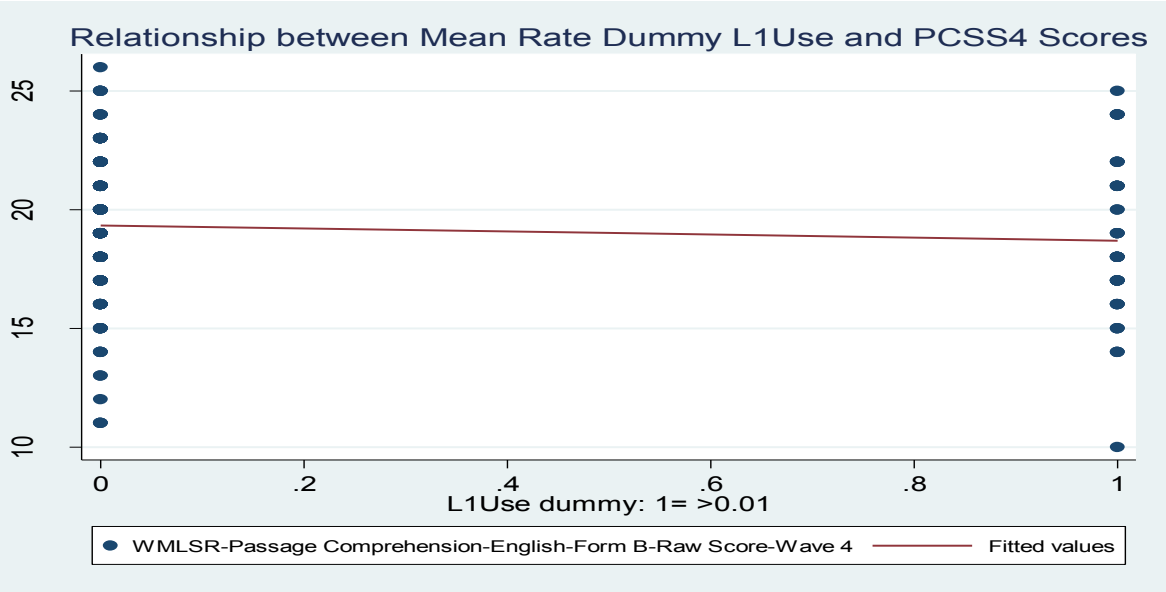


Figure 4.21 *Relationship between Mean Rate Dummy L1Use and PCSS4 Scores*

This review of the descriptive statistics of the classroom-level codes in these language arts lessons suggests that while these codes are not ideal normal distributions, all but L1Use are in the range of normality to use as continuous variables in the following HLM models. Furthermore, these descriptive statistics provide an initial window into various features of instruction observed in these 31 classrooms. Across all three grade levels, teachers in all three grades talked much more than their students, and asked more test questions than uptake questions, eliciting lower-quality cognitive activities like retell and recall. Fourth-grade teachers were significantly more likely to use test questions than fifth-grade teachers, and were significantly more likely to use low-level evaluations as well. All grade levels showed similar low average rates of teacher and student explanations, and little use of Spanish during instruction in any of the classrooms in either district.

Preliminary Analyses: Student-Level Measures

Table 4.8 shows the average scores on all measures of literacy and language included in the final models at the student-level (Level-1). For ease of interpretation of the average student compared to his/her peers, age-equivalent standard scores on the outcome variable (PCSS4) are used in the HLM models. Standard scores describe a student's standing within his/her group, based on a scale with a mean of 100 and a standard deviation of 15. Students in the classrooms that participated in the teaching observations scored, on average, slightly higher on the WMLS-PC in the spring (PCW4) than the fall (Wave 3) PCW3. In age equivalent, standard score terms, students lost a 0.54 standard score over the observation period. This means that, on average, students began the year just below the mean on the WMLS-PC test, and remained below the mean by the spring posttest. The difference in waves 3 and 4, age-equivalent standard scores was not significantly different from zero.

Table 4.8 *Mean (SD) and Range of Literacy Measures used in Level-1 Final Models*

Measure	M (SD)	Min	Max
Outcome			
WMLS-PCSS4	94.41 (10.94)	61.00	127.00
Covariates			
WMLS-PCSS3	94.83 (11.70)	56.00	133.00
WMLS-LWSS3	102.53 (16.61)	9.11	198.00
WMLS-PVSS3	100.43 (14.91)	45.00	151.00
CELF-FS3	36.73 (10.35)	4.00	55.00
CELF-WC3	9.09 (3.30)	1.00	18.00
EB3	36.78 (9.40)	10.00	56.00

Note. N = 236. WMLS-PCSS4 = Woodcock Munoz Language Scales Passage Comprehension standard scores Wave 4 (Spring). WMLS-PCSS3 = Wave 3 (Fall). LWSS3 = WMLS Letter Word ID standard scores Wave 3. PVSS3 = WMLS Picture Vocabulary Wave 3 standard score. CELF-FS3 = Clinical Evaluation of Language Fundamental Formulated Sentences Wave 3 raw score. CELF-WC3 = CELF Word Classes 2 Wave 3 raw score. EB3 = Extract the Base Wave 3 raw score.

Classroom-level and student-level correlations.

Correlations between all covariates included in the final models also provides a window into what patterns and variations in dialogic instruction are seen in upper elementary English language arts instruction (RQ1). Table 4.9 includes all covariates included in the student-level (Level-1) and classroom-level (Level-2) models. Correlations between the student-level language measures were all statistically significant, ranging from a Pearson *r* of .33 to .64 ($p < .001$). Initial correlations of instructional moves (Level-2) indicate that the vast majority were positively correlated with each other. Most correlations were in the low to moderate range. A number of low and moderate correlations are to be expected in classroom discourse data, as it is likely that teachers who tend to use high rates of talk moves likely use many of these talk moves, and teachers with low rates may use less variety. The Bonferroni adjustment was used to account for the multiple comparisons made between 18 covariates. Some notable high correlations were the relationship between test questions (TQ) with low-quality questions ($r = 0.83$, $p < .001$) and low-quality teacher evaluations (TEvalL) with low-quality questions ($r = 0.83$, $p < .001$), which may suggest some consistency in emphasis of language and tasks in classrooms. High-quality questions were much more strongly correlated with uptake questions (Up) than with TQs ($r = 0.80$, $p < .001$ and $r = 0.30$, $p < .001$, respectively). However, there was a high correlation between Up and low-quality questions as well ($r = 0.75$, $p < .001$), suggesting that Up showed more of a range of question quality than TQ. Interestingly, student explanations (StExp) were (moderately) correlated with uptake questions ($r = 0.44$, $p < .001$) and high teacher evaluations (TEvalH; $r = 0.58$, $p < .001$), suggesting a possible dialogic relationship between these two teacher talk moves and increases in student talk. StExp showed no correlation with low-quality questions and TQ, suggesting that test questions asking students to recall or repeat known

information did not elicit extended student answers. First-language use (L1Use), rare in all classrooms, showed no significant correlations with any Level-1 or Level-2 measures.

Table 4.9 *Correlations across Level-1 and Level-2 Measures*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.PCSS4	--																
2.PCSS3	0.64*	--															
3.LWSS 3	0.56*	0.53*	--														
4.PVSS3	0.55*	0.63*	0.43*	--													
5.FS3	0.51*	0.55*	0.35*	0.55*	--												
6.WC3	0.40*	0.40*	0.33*	0.42*	0.59*	--											
7.EB3	0.45*	0.51*	0.54*	0.37*	0.51*	0.55*	--										
8.farms	-0.28+	-0.27+	-0.14	-0.37*	-0.28+	-0.21	-0.11	--									
9.firstlang	-0.19	-0.24~	-0.11	-0.46*	-0.32*	-0.15	-0.11	0.37*	--								
10.TQ	0.15	0.13	0.03	0.03	-0.02	-0.04	0.03	-0.03	-0.11	--							
11. Up	0.08	0.06	-0.07	0.04	<.01	0.13	0.10	0.01	<-.01	0.40*	--						
12.High Qual	0.01	<-.01	-0.05	-0.01	-0.03	0.04	0.04	0.04	0.07	0.30*	0.80*	--					
13.Low Qual	0.15	0.13	-0.05	0.05	<-.01	0.06	0.05	-0.06	-0.10	0.83*	0.75*	0.53*	--				
14.TEval H	0.06	0.07	-0.06	0.06	0.13	0.22	0.14	-0.13	-0.09	0.35*	0.62*	0.62*	0.62*	--			
15.TEval L	0.21	0.13	0.01	0.12	0.09	0.19	0.15	-0.15	-0.10	0.72*	0.60*	0.36*	0.83*	0.54*	--		
16.TExp	0.16	0.11	0.04	0.13	0.12	0.05	-0.01	-0.19	-0.10	0.30*	0.07	0.31*	0.32*	0.53*	0.22	--	
17.StExp	-0.01	0.02	-0.08	0.10	0.21	0.34*	0.22	-0.17	-0.15	-0.22	0.44*	0.22	0.19	0.58*	0.25+	0.11	--

18.UseL																
1	-0.10	<-.01	-0.17	-0.08	0.03	0.01	<-.01	0.01	0.14	-0.03	0.12	0.14	0.08	-0.04	-0.02	-0.05	0.13

~ $p \leq .05$. + $p \leq .01$. *$p \leq .001$.

Modeling the Relationships between Instructional Talk Features and Student Reading Achievement

To answer the question of which indicators of dialogic instruction predict students' reading comprehension scores (RQ 1a) on the Woodcock Muñoz Language Survey- Revised, Passage Comprehension subtest (WMLS-PC; Woodcock et al., 2005), hierarchical linear modeling (HLM) was used to generate estimates for the final model. Stata software (version 12.1) was used for this stage of the analyses.

The HLM models were built in a hierarchical method, and results from each model are reported in Table 4.10. In all models, centered scores were used for all measures, except the two demographic, binary measures of free and reduced lunch status (farms) and first language status (firstlang). Unstandardized coefficients are presented in the final model.

To control for students' initial language acumen, vocabulary and reading measures were included in first level of the models. The fall wave (Wave 3) of the WMLS-Passage Comprehension standard score (PCSS3) was included to control for students' initial reading comprehension levels. Students' initial breadth of vocabulary knowledge (PVSS3; Woodcock et al., 2005), word identification (WMLS-LWSS3), and three other word knowledge measures were included in the Level-1 model. Students' syntactic (CELF-FS3; Semel et al., 2003), semantic (CELF-WC3), and morphological knowledge (EB3; Anglin et al., 1993; August et al., 2001; Carlisle, 1988) were also included in the student-level models. Raw scores on these latter three measures were used. Table 4.8 (above) presents the means and standard deviations of these covariates in the sample. The means for all measures were from students who completed both

waves of the outcome variable (PCSS4; student $n = 236$) and were in the classrooms that were observed for instructional features.

Table 4.10 presents the final models calculating the relationships of dialogic instructional features and reading comprehension. Model 1 is an intercept-only model which partitions variance by accounting for the nesting of students within classrooms. The grand mean of 94.59 ($SE = 1.10$) indicates the average PCSS4 score for students when accounting for nesting within classrooms. A significant likelihood ratio test statistic of 16.63 ($p < .001$) confirms that there is between-classroom variation.

Model 2 provides the parameter estimate for students' initial reading comprehension scores (PCSS3), showing a significant and predictive relationship between pre- and post-test reading comprehension scores ($\beta_{1j} = .58$, $p < .001$). The pseudo-R^2 indicates that students' initial scores on the PCSS3 accounted for 32% of the student-level variance in the outcome. The other two pretest WMLS-R measures were added into Model 3 to account for students' fall vocabulary and word recognition scores. The pseudo-R^2 indicated that about 11% of additional outcome variability was attributable to these covariates.

When the three pretest vocabulary depth measures (FS3, WC3, EB3) were added to Model 4, only word recognition and syntactic/semantic achievement were significantly related to students' PCSS4 achievement. The other measures did not meet the .05 significance level and were dropped from subsequent models. The parameter estimates for this more parsimonious model are reported in Model 5 of Table 4.10. These additional student measures accounted for another 1% of variability in the outcome. This model of reading indicates that, accounting for students grouped in classes, students' fall reading comprehension, letter word identification skill,

and expressive syntactic and semantic knowledge contribute to explaining 48% of the variance in student reading comprehension scores in this sample.

Two demographic variables were entered in Model 6. Socioeconomic status (farms) was negatively and significantly related to PCSS4 across the student sample at $\beta_{7j} = -2.48$, $p < .01$. Students' first language was not related to their achievement on the PCSS4, but was kept in the final model to address the focus of the research questions on bilingual students. Thus, the final baseline, Level-1 model, accounting for variance in students' scores at the student level, included only those measures that were significant with an alpha level of .05 or below, except for first language. This model's estimate of the total variance between classrooms was 7.85. The pseudo-R^2 calculation indicated that these five covariates accounted for 49% of the variation of student outcome scores. In the subsequent final model, these variables were interpreted to control for students' initial language and reading achievement, socioeconomic, and first language status.

Model 7 (Table 4.10) is the final HLM model which includes parameter estimates at both levels in the data. Literature on dialogic instruction has nothing to say about an order to entering instructional predictors into such a model. Thus, the nine Level-2 predictors used to code instruction in this study were entered at once into the final model in order to examine their relative importance as predictors of student reading scores. Uptake questions, low-quality evaluations, and teacher explanations were positive predictors of reading outcomes and significant main effects in this final model. Student explanations and high-quality questions were significant negative predictors of student outcomes.

Teacher explanations were the strongest predictors ($\gamma_{07}(28.46); p < .01$). When using this unstandardized parameter in the final model equation, an average student in a classroom with an average teacher explanation rate would have a predicted score on the PCSS4 of 95.68. A half

standard deviation increase in the rate of teacher explanations would predict an average student's score of 96.82, two standard score points higher than the sample's mean on the outcome measure. The coefficient estimated for uptake questions ($\gamma_{02}(25.13); p < .05$) indicates that if an average student attended a classroom with a half standard deviation increase in the uptake rate, the student's predicted score would be 97.95 on the PCSS4, a 3.54 increase of standard score points. Similarly, if an average student experienced instruction with a half standard deviation higher rate in low-quality teacher evaluations (TEvalL ($\gamma_{06}(10.18); p < .05$), the student reading scores would be predicted to be 96.80 standard score points, or about 2 standard score points higher on the outcome measure from the sample mean.

Student explanations predicted lower reading scores. An average student in the average rate of student explanation classroom would be predicted to score 95.69 points on the PCSS4, or 94.26 if the average student experienced instruction with a half standard deviation higher rate of student explanations. Teacher high-quality questions (asking students to infer, predict, synthesize or evaluate content) were also shown to be significant predictors of student achievement, and also indicated a negative relationship: specifically, for a change of a half standard deviation in the rate of high-quality questions, the average student would be predicted to score 94.18, about the sample mean on the PCSS4. Conversely, if this student were to experience instruction characterized by a half standard deviation drop in the rate of high-quality questions, s/he would score 97.11 on the outcome measure, just over two and a half points higher on the age-equivalent, standard score scale. A discussion of the possible reasons for this counter-intuitive finding is included in Chapter Six.

Calculation of a pseudo-R^2 statistic showed that these Level-2 covariates account for about 82% of the outcome variable. Thus, the study's hypotheses that the features of dialogic

instruction would be predictive of students' reading comprehension scores were only partially true: while uptake questions do predict students reading scores in this sample, the other instructional variables of high-quality questions, low-quality teacher evaluations, and teacher and student explanations disconfirmed the hypotheses that features of dialogic instruction are related to higher reading outcomes.

Table 4.10 *Parameter Estimates (SE) and Model Fit for Hierarchical Models 1-7*

	Models						
	1	2	3	4	5	6	7
Fixed Effects							
Intercept, β_{0j}	94.59*** (1.10)	94.46*** (.67)	94.45*** (.63)	94.39*** (.71)	94.41*** (.70)	95.85*** (1.09)	95.97*** (.99)
WMLS Passage Comp. W3, β_{1j}	.58*** (.05)	.33*** (.06)	.24*** (.06)	.29*** (.06)	.28*** (.06)	.27*** (.06)	
WMLS Letter Word W3, β_{2j}		.20*** (.04)	.1*** (.04)	.21*** (.03)	.21*** (.03)	.20*** (.03)	
WMLS Picture Vocab. W3, β_{3j}		.13** (.04)	.07 (.04)				
CELF Formulated Sentences W3, β_{4j}			.16** (.07)	.24*** (.06)	.23*** (.06)	.24*** (.06)	
CELF Word Classes W3, β_{5j}			.26 (.20)				
Extract the Base W3, β_{6j}			.04 (.07)				
Farms, β_{7j}					-2.48* (1.23)	-2.20 (1.21)	
First Language, β_{8j}					.77 (1.09)	.42 (1.10)	
Test Questions, γ_{01}						-9.82 (7.93)	
Uptake Questions, γ_{02}						25.13* (10.83)	

High Quality Questions, γ_{03}							-16.69* (7.95)
Low Quality Questions, γ_{04}							-3.41 (7.01)
High Evaluations, γ_{05}							-2.17 (6.36)
Low Evaluations, γ_{06}							10.18* (4.57)
Teacher Explanations, γ_{07}							28.46** (10.32)
Student Explanations, γ_{08}							-23.73** (9.36)
L1 Use, γ_{09}							-0.44 (1.39)
Random Effects							
Level 1 residual, σ^2	97.11 (9.63)	65.03 (6.38)	53.60 (5.26)	48.85 (4.86)	50.25 (4.95)	49.44 (4.87)	49.77 (4.91)
Level 2 variance, τ_{00}	23.70 (9.95)	5.02 (3.45)	5.23 (3.17)	8.49 (4.18)	8.10 (3.85)	7.85 (3.76)	1.44 (2.15)
Model Fit							
Deviance	1781.32	1669.09	1626.23	1612.96	1618.38	1614.29	1598.03
ICC	.20	.07	.89	.15	.14	.14	.03
Δ Level-1 pseudo R^2		.32	.11	.05	.01	.01	
Level-1 pseudo R^2		.33	.45	.50	.48	.49	
Level-2 pseudo R^2							.82

Note. R^2 and ΔR^2 left blank where not applicable. Unstandardized parameter estimates reported.
* $p < .05$, ** $p < .01$, *** $p < .001$.

The second research question addressed by these quantitative analyses on whether bilingual and monolingual students are differentially impacted by exposure to high levels of dialogic instruction can be answered by Model 7. The effects of the five significant instructional

talk moves were main effects across the sample, regardless of first language status. Thus, these results are predictive of both bilingual and monolingual student scores.

Chapter 5 – Results: Qualitative Strand

The quantitative analyses reported in Chapter Four showed that higher rates of uptake questions, low-level teacher evaluations, and teacher explanations used in a classroom were associated with positive changes in students' reading comprehension scores. Conversely, high rates of high-quality questions and student explanations were associated with negative changes in reading outcomes. All five findings showed main effects across all grades, SES status, and for both bilingual and monolingual students. As a convergent mixed method design study, an exploration of the patterns and variations of each finding is presented qualitatively in this chapter to answer the second research questions driving the study: 1) *What patterns and variations in the significant findings of dialogic instruction from RQ1a are seen across the classrooms,* and 2) *What variations of significant features of dialogic instruction are seen within classrooms*?

Each talk move (or predictor in the quantitative analyses) is analyzed here as a case in its "real-life context" of classroom instruction (Yin, 2009, p. 18). We consider each talk move a case because it is a "phenomenon … occurring in a bounded context" (Miles & Huberman, 1994, p. 25), each talk move being bounded by its lesson, and by its particularity compared to other moves during instructional talk in this sample of upper-elementary classrooms. As Stake (2006) argues, a case is "a noun, a thing, an entity" rather than a verb or a function, and we understand a case by "experiencing the activity of the case as it occurs in its contexts and in its particular situation," to find out "how the case gets things done" (p. 1-2). By presenting uptake questions, for instance, as a singular case, the quality of this question type can be explored to better understand how these questions might be precipitated by certain talk moves, or what patterns of talk might follow these questions. By looking more inductively at each of the five moves as they function within and across language arts lessons, the intent of this analysis is to consider how

they function within patterns of talk to, at least partially, support students' reading comprehension achievement. Therefore, patterns and themes (Coffey & Atkinson, 1996) are identified in each case. A multiple case study analysis, where one analyzes how each case is related to other cases (Stake, 2006), is beyond the scope of the current study but a topic important for future analyses. Since the quantitative analyses treat each talk move as an individual predictor of reading comprehension, the five talk moves will be treated as unique in these qualitative analyses as well.

Until this chapter, the analyses have assumed that the rate of dialogic features like uptake questions represent sufficient instructional quality to impact student reading scores on the Woodcock Muñoz Language Scales – Revised Passage Comprehension reading test (PCSS4; Woodcock et al., 2005). In the following analyses, while still drawing on frequencies of codes as a proxy for quality, the analytical emphasis is on how these five talk moves represent instructional quality in and of themselves. By organizing these analyses as cases, we hope to capitalize on the strength of case studies are having a particularistic focus and thus helpful for "practical problems… arising from everyday practice" (Merriam, 1998, p. 29), as well as to move beyond previous correlational analyses that can assume an "X yielding Y" relationship between instruction and student achievement, critiqued as simplistic for instructional discourse studies (Nystrand, 2006).

The transcription data were analyzed using Atlas ti 6.2 qualitative software (Scientific Software Development, 2011). To identify themes and patterns of the case, the first analytic step was to read through all instances of the case across all transcripts to identify and categorize the content of the talk move. For example, high-quality questions elicited inferences, evaluations, and a few analyses from students. General themes were identified and refined from these initial

codes to identify the most common patterns in the content of the talk moves, to condense data into analyzable and identifiable units (Coffey & Atkinson, 1996; Miles & Huberman, 1994). For instance, high-quality questions were highly scaffolded to narrow the possible responses to questions. The second analytic step was a search for the patterns exhibited by the talk move in the transcripts: what other talk moves accompanied it, or which talk moves preceded or followed it. Using the Atlas ti "query" function, this focus allowed for a better understanding of the case's function and place in literacy learning activities.

A third analysis addressed both RQs to identify variations of each talk move within classrooms as well as across classrooms. To do this, patterns of the talk move were identified and thematically categorized by looking at their functions in high-rate and low-rate classrooms. Since the rates of the five talk moves were found predictive of reading scores, differences in classrooms characterized by extensive use and classrooms that used the move infrequently were explored. The distributions of the talk showed slightly positively skewed distributions (Chapter 4 Results), with a few high rates of the move and more low rates, the median is a more appropriate measure of central tendency (Shavelson, 1996), used to identify high- and low-rate classrooms. Therefore, I grouped classrooms into "high" and "low" talk move classrooms by their inclusion in the 25^{th} and 75^{th} percentiles to get a better sense of the discrepancies of pedagogical talk between classrooms with different talk move tendencies. This third component of the analysis allowed comparisons across groups of classrooms characterized by the dichotomous variable (high or low rate) and also allowed comparisons across lessons by the same teacher. These within classroom analyses are embedded throughout each case study.

These three analytic steps were undertaken to provide a particularist picture of each case (Merriam, 1998; Stake, 2006). Discourse data included in this chapter are cited according to

Atlas ti nomenclature: the primary document number (P12), the number of the quotation, or code, applied to the data (P12:48), and the date of the observation (P12:48, 2/15/11).

To situate these five significant talk moves in the classroom observation data, the first section briefly summarizes the descriptive statistics of four global features of instruction across all 31 classrooms: teacher or student management, pedagogical materials, talk genres, and language arts topics addressed in the lessons (codes described in Chapter Three). Following this section are each of the five talk moves presented as cases, in which they are defined, exemplified, and their patterns explored across classrooms and within classrooms.

Global Features of Instruction: General Trends across the Lessons

As described in Chapter Three in the "First Stage of Qualitative Analyses" section, each of the 88 transcripts of language arts lessons was coded at a global and a more proximal level of instruction. Global features of instruction included identifying the locus of lesson management, instructional materials used during instruction, lesson topics, and general genres of talk through which the instruction occurred. Proximal features of instruction were the nine particular talk moves used as predictors in the quantitative analyses. A summary of the trends of the global features are presented here to provide context for the five cases of talk moves that follow.

Transcripts were coded for global features of instruction by coding instructional episodes. Instructional episodes were coherent learning activities that functioned to meet a particular objective or goal (Nystrand et al., 2003), segments of lessons that were "coherent internally" through their focus on a particular learning activity (Linnell, 1998, p. xiv). Most lessons were organized around one to three instructional episodes; a common lesson would begin with a review of the plot of a chapter book the class was reading, followed by a read aloud, where the teacher typically addressed new vocabulary and/or a particular reading strategy and checked

students' comprehension of the unfolding story, and then led into another learning activity, like a short writing project that incorporated a graphic organizer. Each episode was coded for lesson management, instructional materials, talk genres, and the field (topic) of the lesson.

In the majority of the classes, and during the majority of the lessons, teachers were responsible for "focusing attention on learning" (Connor et al., 2007, p. 209) rather than students, who, for example, would tend to manage their work in small group or independent seat work. Teachers were twice as likely as students to drive the progress of instruction by controlling turn-taking to advance the lesson for student learning across all three grade levels.

Table 5.1 *Means (SD) of Teacher- and Student-Managed Instructional Episodes by Grade*

	Teacher-Managed Episodes		Student-Managed Episodes	
Grade	Mean (SD)	N	Mean (SD)	N
Grade 3	2.11 (0.96)	12	0.88 (0.47)	12
Grade 4	2.48 (0.72)	11	1.21 (0.56)	11
Grade 5	2.33 (1.17)	8	1.15 (0.71)	8

Two-tailed, independent means t-tests revealed no significant differences between any of the grades in either type of lesson management.

Instructional materials were first coded by their name used in the transcripts or field notes (e.g. "Open your anthology to page 56"), and later refined to 10 categories of materials used across the sample. Two or three types of instructional materials were sometimes used concurrently; for instance, a few classes used graphic organizers with student texts and practice tests to organize students' ideas. Unabridged narrative texts and testing materials were the most frequently used across the lessons, followed by basal texts (literature anthologies or ELA textbooks) and student-generated texts. Unabridged narrative texts were also used more frequently during an entire lesson, over multiple instructional episodes, indicating their use in a

variety of learning activities with differing goals. Most of the other materials were not used or used only once during a lesson. Appendix B presents a table of the frequencies of each material used across the 88 transcripts.

Instructional episodes in each transcript were also coded for talk genres to identify the general patterns of discourse used to deliver literacy lessons. The term 'genre' here is drawn from Systemic Functional Linguistics (SFL) theory, where a genre is a "staged, purposive activity" structured in ways to serve instructional goals (Christie, 2002, p. 21). These codes identified the patterns of pedagogical talk established by the teacher to deliver each instructional episode. These codes were derived inductively, grounded in the transcript data to identify and describe the patterns of instructional talk and were refined to seven categories, including question and answer, small-group and whole-class text talk, and direct instruction. The question and answer talk genre (Q & A) was the most frequently used talk genre across the classrooms, followed by read alouds and individual work activities. The Q & A pattern of talk differed from the whole class text talk code in instructional focus: in the Q & A genre, the instructional focus was on answering a series of questions, while text talks were characterized by more extended talk about text with some questions building to more depth than questions that move to the next topic. The Q & A genre, in conjunction with the high levels of teacher-managed instruction, are suggestive of a predominant IRE pattern of instructional talk across the grades. Appendix B includes the frequencies of each of the seven talk genres. It is important to note that the averages shown in Appendix B do not account for the length of transcript or number of instructional episodes per transcript. These frequencies are based simply on the number of talk genres coded in each instructional episode, so it cannot be concluded that students spent more absolute time in

the Q&A talk genre, though this is probably most likely in most classes given the high number of occurrences of its use.

The fourth group of global codes, fields (topics) of instruction, were derived from the transcript data and refined into 10 categories that summarized the language arts topics addressed in these classrooms. The most frequent focus of instruction in these language arts classes was students' text comprehension of their reading materials across all three grades. Vocabulary, writing, and to a lesser extent, reading strategies, were also addressed in many instructional episodes. The modeling of exemplary work and the development of students' background knowledge for a learning activity were two of the least frequently observed topics of instruction. The frequencies of each language arts field addressed in the lessons are summarized in Appendix B.

Finally, presented early in Chapter Four, but worth noting again here, was the preponderance of teacher talk and a question and answer pattern of instructional talk across all three grades and classrooms. Across the sample, the mean rate of total talk was 93.20 words spoken per minute (*SD* = 36.31), while teacher talk (total words minus student words in each transcript) averaged 77.86 words spoken per minute (*SD* = 30.33). Therefore, teacher talk represented 84% of the talk during these lessons. Within this talk, teachers across the sample posed an average of one question per minute (*M* = 1.07, *SD* = .45). Table 5.2 presents these overall talk rates disaggregated by grade level.

Table 5.2 *Mean (SD) Rates of Talk, Teacher Talk and Questioning by Grade*

	Overall Rate of Talk		Teacher Talk Rate		Teacher Question Rate	
Grade	Mean (SD)	N	Mean (SD)	N	Mean (SD)	N
Grade 3	87.50 (38.52)	33	77.23 (32.81)	33	0.98 (0.47)	33
Grade 4	96.67 (32.64)	33	79.76 (26.97)	33	1.32 (0.66)	33

Grade 5	96.56 (38.69)	22	75.96 (32.42)	22	0.85 (0.40)	22

Note. N = number of transcripts; Overall and teacher rate of talk = Audible words per minute; Teacher question rate = Audible instructional questions per minute

As Table 5.2 indicates, there were high levels of teacher talk across the grade levels, though with such high standard deviations, it is clear that there was a lot of variation among classes. Average teacher questioning rates were not as uniform across grades, with fifth-grade teachers posing fewer questions, on average, than the other two grades. Uptake questions were posed more frequently than other types of questions in grades four and five, while third-grade teachers posed test questions most frequently. The general trend was that all grades were characterized by high levels of teacher talk and questioning, and these high levels of teacher talk were commensurate with high levels of teacher-managed instruction.

In summary, the global features of instruction presented above were coded at the episodic level of instruction and do not account for the differing lengths of each lesson. However, frequency statistics indicate some general trends of the instruction in this sample. There was a prevalence of teacher-managed instruction and a preponderance of teacher talk, with teachers' speaking for about 84% of the observed lessons, on average, across the three grade levels. While there were wide variations of use of instructional materials, unabridged narrative texts were the most frequently used materials. Teacher-led question and answer talk patterns were the most frequent pedagogical discourse pattern, and a focus on the comprehension of texts was by far the most frequent topic addressed in these lessons.

Utterance Features of Instruction: Patterns and Variations in Talk Moves Across and Within Classrooms

There were a total of 21,176 codes applied to 88 transcripts, summarized in the coding scheme (Appendix A). Low-quality questions and low-quality teacher evaluations were the most

frequently applied code to the data, followed by high-quality teacher evaluations and the two kinds of questions kept for final analysis, uptake and test questions (as described in Chapter Three and Appendix A, there were four types of questions coded, but only test and uptake questions had high enough coding reliability to be included in these analyses).

Teacher explanations, uptake questions, and low-quality evaluations, in descending order of strength, were the three positive predictors of reading achievement. Student explanations followed by high-quality questions were predictive of lower student reading comprehension scores in the final models. Averages and ranges for each of the four talk moves are presented in Table 5.3. These statistics are the basis for some of the following qualitative analyses on each of the five talk moves. The five talk moves as cases are presented in order of their strength of relationships with the reading outcome measure, starting with those cases that predicted higher outcomes. This chapter concludes with the two of five cases that predicted lower reading comprehension scores.

Table 5.3 *Descriptive Statistics for the Five Significant Talk Moves from HLM Final Models*

	Raw Counts						Rates per minute					
Code	Mean (SD)	Min	Max	Median	Q25	Q75	Mean (SD)	Min	Max	Median	Q25	Q75
Positive Relationship with Reading Comprehension												
Low-quality Evaluations	30.53 (20.00)	1	104	25	17	41	0.56 (0.22)	0.16	1.35	0.47	0.37	0.70
Uptake	17.55 (13.84)	0	80	14	9	24	0.32 (0.19)	0.09	0.99	0.27	0.19	0.40
Teacher Explanations	7.23 (5.61)	0	33	6	3	10	0.12 (0.08)	0.20	0.33	0.11	0.07	0.18
Negative Relationship with Reading Comprehension												
High-quality Questions	17.81 (13.02)	0	80	15	9	23.5	0.33 (0.18)	0.05	1.01	0.28	0.23	0.41
Student Explanations	9.58 (9.39)	0	34	6.5	2	14.5	0.16 (0.12)	0.00	0.44	0.11	0.07	0.15

Notes: Raw counts (n=88); Rates per minute (n=236); Q2 = 25th percentile/quartile score, Q75 = 75th percentile score. Teacher explanations were the strongest positive predictors of reading, followed by uptake, and then low-quality evaluations. Student explanations were the strongest negative predictors of reading, followed by high-quality questions.

Case 1: Teacher Explanations.

Teacher explanations were hypothesized not to be indicators of dialogic instruction but may work to support students' reading comprehension indirectly through vocabulary knowledge (Aukrust, 2007). While this finding cannot be corroborated by the current analysis, teacher explanations in this sample were shown to be the strongest predictor for reading comprehension scores in the average classroom. A half standard deviation increase in the average rate of teacher explanations predicts an average student score of two points (standardized scores) higher on the outcome measure. These extended turns of talk consisted of two or more lines of transcript and were usually characterized by two or more ideas. These explanations were 'stand alone' instructional utterances that included no questions and were not in response to a student's question or input. These codes were not as frequently applied to the transcript data as the following two codes that also predicted higher scores; there were a total of 638 utterances codes as teacher explanations, accounting for 3% of the codes applied to the 88 transcripts. Teacher explanations average about seven per hour of instruction across the 31 classrooms, but this average varied widely across the sample. Teacher 4 used the highest number of explanations at an average rate of 0.33 per minute. Her third-grade colleagues used explanations more frequently ($M = 0.14$, $SD = 0.10$) than the other grades, and fifth-grade teachers the least ($M = 0.10$, $SD = 0.08$). Teacher explanations ranged from two sentences with two ideas ("There is a national popcorn day, believe it or not. It was last Friday") to 19 sentences of uninterrupted speech.

Across both sites and across all grade levels, teacher explanations, in descending order of frequency, functioned to 1) provide definitions and descriptions to support understanding of a given text topic, 2) review and summarize previous content, 3) offer "meta" comments of class work, and 4) model reading strategies.

Teacher explanations as descriptions.

The most frequent function of teacher explanations was to define concepts and vocabulary, employing examples in many cases. The following examples compiled on Table 5.4 are typical explanations that functioned to support content understanding in these language arts classes.

Table 5.4 *Typical Examples of Teacher Explanations as Descriptions of Content*

Teacher	**Speaker**	**Discourse and Teacher Explanations** (in bold)
50	Teacher:	**When something's stale it is no longer fresh. Who has ever maybe had a piece of bread that was left in the package for a long time. It was kind of hard, it didn't taste as good. Stale bread**. [P8:30, 2/8/11]
4	Teacher:	**Yesterday my sister and I went up in the attic.** You know what you have up in the attic. What do you have up in the attic?
	Students:	[*overlapping voices*] Junk. Books.
	Teacher:	Yeah, **junk and books and all that other stuff. So you know what we found up there? My father's hammock from the Navy!**
	Student:	What?
	Teacher:	**It's a big piece of rope, but it's like a cushion** [*unintelligible*] **hammock and that's what he used to sleep on, on the ship when he was in the Navy. So that's a hammock.**
	Student:	Miss [Teacher 4], I have something. My dad used to play hockey and we went up in my attic and found all his hockey stuff. [P28:19, 1/27/11]
62	Teacher:	Thank you. **The horizon is the line far away where the earth and the sky seem to meet. OK? You look far out, where you can see the horizon. So here's a really nice picture of the horizon. Sometimes when you look at the horizon, the earth looks flat. So you can see right here is the horizon, when the earth and the sky seem to meet.**
	Student:	It looks the movie [*unintelligible*] when there were all these pink flowers on the horizon.
	Teacher:	OK, excellent. Good connection. So you've seen the horizon before.
	Student:	Marcelo.
	Teacher:	Like the [*unintelligible*] is like a hill, I think. No, these are just flowers in a field, and this is the sky. And the horizon is where the earth and the sky seem to meet. OK? Where they seem to meet. Like when you look out, it's hard around here, because there's a lot of houses. But if you're in like an open field area, or like the ocean, where you like the sky and the water seem to meet, you look out over the water, it's easier to see as well. [P10:10, 12/2/10]
53	Student:	[*reading text aloud to class*]

Teacher: Good. **Here we have another example of how Karana is knowledgeable about the island and the water. It says that she paddled from the back of the canoe.** Why did she paddle from the back of the canoe? [P46:104, 3/1/11]

Note. Words in bold represent coded instances of teacher explanations

Many teacher explanations were used to define concepts and provide examples, exemplified by excerpts from Teachers 50, 4, and 62. Teacher 4 provided an example of typical things found in an attic, and this precipitated a student's own extended talk move. The teacher's descriptions seemed to function as an invitation to share personal examples. Teacher 62 provided a lengthy description of a horizon, referencing an accompanying visual. When a student contributes an off-target response about a hill, Teacher 62 provided what was coded as a high-quality evaluation (a follow-up move), rather than another "stand along" explanation, because this turn was in direct response to the student's (erroneous) contribution. In the excerpt from Teacher 53, her example built to a question attending to student comprehension of the text.

Teacher explanations were not commonly precipitated or followed by any type of question in these data, with only one identifiable pattern involving teacher explanations. Explanations tended to be used independently of the pervasive questioning pattern in most classrooms, and most other talk moves coded for in this study. In about five percent of teacher explanation codes, they followed low-quality evaluations (in italics), as the following example shows.

Teacher 82: Now, here's what's interesting, almost all of these answers could work if you just plugged it the spot. Right? The forest fire climbed the mountain side?
Student: No.
Student: Yes.
Teacher: *Yes.* **That's using personification, but think about it. When something catches on fire, notice how it just kind of keeps on creeping and keeps on burning things up? Because a fire can start down here, but it just keeps consuming. So it keeps moving. So that makes sense. The forest fire covered the mountainside.** Doesn't that make sense?

Student: Yes.
Teacher: Yeah, that makes sense. [P71:57, 12/8/10]

The students' answer to the teacher's question ("The forest fire climbed the mountainside?") was met with a simple affirmation ("yes"), repeating a student's answer. Teacher 82 then offered an extended explanation for why the personification of fire can make sense in the cloze question (previously addressed, not included here). In this example, the explanation work in the same way as teacher evaluations, occupying the 'E' spot in the IRE pattern, providing information, but not in a reciprocal fashion where the teacher responds to a student's contribution by correcting or extending it. Teacher explanations work more like direct instruction, telling students what content is important to understand.

Teacher explanations as summaries.

A second theme identified in this case study of explanations was the use of explanations to summarize content and previous work. Many times these were provided by teachers after a read aloud. In these instances, the teacher did not pose a question about the text at first, but instead provided students with a direct rephrasing of the key plot events just covered in the reading. Two typical examples are presented in Table 5.5.

Table 5.5 *Typical Examples of Teacher Explanations as Summaries*

Teacher	Speaker	Discourse and Teacher Explanation as Summary
27	Teacher:	We talked about how character traits are how we describe a person. And the choices for the character traits should be things that carry over into all parts of your life. So if somebody's patient in one respect and they're not in another, you couldn't really call it a true character trait. So you can use some of the ones that you used in your actual response, but we want to find it right in the text. OK. So looking back here – (P1:29, 1/28,11)

18	Teacher:	Now remember I told you how we fight over the sweater and it's not really attractive, but it's so, doesn't it look like it's torn? OK? And it's more special than it is anything. And this is something that I hope to have, even though it's falling apart, because it's old. This is what I would bring [with me if I moved]. So I would illustrate a picture of this sweater and I would explain what it means to me. OK? Does that make sense? So although to you, you might think, my goodness, why would [I] bring that god awful ugly sweater with [me] if [I] left the country [*overlapping voices*]. Why would I? Why would I bring this sweater? Why would this be important for me to bring, Jaoquin? (P78:74, 1/24/11)

Teacher summaries of texts or previous work were not followed by consistent patterns of talk moves, other than teacher questions "book ending" the explanations, because most of the lessons were conducted using a Q & A genre of instructional discourse. Teacher 18, above, exemplifies this pattern, finishing her explanation with four rapid questions.

Teacher explanations as "meta" commentary.

A third pattern to teacher explanations was the use of these talk moves to focus students' attention on the work of the class. These explanations attended to both content and how the content was accomplished. Three examples are presented in Table 5.6.

Table 5.6 *Typical Examples of Teacher Explanations as Commentary*

Teacher	**Speaker**	**Discourse and Teacher Explanation as Commentary** (in bold)
27	Student:	Gordie runs away and Doug goes to get him back before he gets in trouble.
	Teacher:	Good. **Now, remember we had some questions about Gordie's age, was Gordie younger or older. We still have the question. We're still not sure about that, but some of you made inferences that he was older, because he wouldn't necessarily have run away, to really run away, yes.** [P3:13, 3/3/11]
23	Teacher:	All right. What I noticed—**one of the things I noticed today, not only when I was walking around, but sharing, is in our writer's notebooks, we're starting to lose a little bit of our organization. People having to shuffle through the pages to find the right page,**

		especially when you're getting ready to share, it makes us lose the
	Student:	**story. We lose—it's called a continuity. A flow.**
	Teacher:	Continuity?
		Continuity. It's how things flow together. If you're in the middle of reading a sentence and you stop and you have to look around and then you keep reading, everyone's forgotten what the sentence was all about…. [P76:63, 2/9/11]
23	Teacher:	**Let's see what Tanisha did. See what Tanisha did. Tanisha made some changes. She—in the book that I used this in color, but in here it's not, so I'm going to add some color for you. She added the word and. I fell and I got hurt, instead of I fell. I got hurt. Is that a good addition to a story, to add that and make more sense?** (P76:3, 2/9/11)

Note. Words in bold represent coded instances of teacher explanations

Explanations functioned as 'meta' comments to provide explicit attention to current or previous work. Teachers 27 and 23 both explained content through explicit focus on the learning tasks of the class, and how these supported the class work. Teacher 27 offered a brief summary of a question the class had about a text, and reminded students of an inference made to support a theory about the text ("Some of you made inferences that he was older, because he wouldn't necessarily have run away"). Teacher 23 explained that students' organization of their notebooks was a key support for sharing their work, and at the same time, explained a tier 2 vocabulary word ("continuity") by contextualizing it within his explanation how to read a sentence so the audience comprehends it. In a second example from the same class, he uses a student's work to focus the class's attention on how one might revise their writing.

Teacher explanations as models.

The fourth and final theme apparent in this case of a predictive talk move of student reading comprehension was the use of explanations as models for reading strategies. These explanations of strategies were mainly focused on test taking in lessons, and more frequently

models of reading informational texts (the most frequently used pedagogical material across these classes was narrative text). Two examples of strategy descriptions are shown in Table 5.7.

Table 5.7 *Typical Examples of Teacher Explanations as Strategy Models*

Teacher	**Speaker**	**Discourse and Teacher Explanation as Model** (in bold)
63	Teacher:	Now, close your eyes. I want you to visualize right now. Close them. **Visualize this. Let's pretend it's the same time of day, tomorrow morning, and you come to school and you are the only person here. There are no teachers. There are no kids. There is no office staff. There is no principal. You come in. The school is unlocked and all the lights are off. There are no computers on. Nothing is happening. You are the only one here. … You walk into a classroom and you see chairs turned over, desks are on the floor and there's not one single person here. It's dark. The school is deserted. There is no one here.** Now, open up. **Imagine that same thing happened to the Maya culture. The Maya civilization, okay? When the Maya people were living there, all of a sudden their civilization became deserted.** [P21:89, 11/30/10]
64	Teacher:	Boys and girls, here's the way I would look at this question. **If I'm trying to decide what the main idea is, I go back to my map and I see what's the main problem in this story. The moose is drinking all the water and the solution is the fly gets rid of that moose. The main idea is that a fly gets rid of a troublesome moose. That's the main thing that happens in the story.** [P34:90, 1/14/11]

Note. Words in bold represent coded instances of teacher explanations

These examples of teacher explanations providing strategy instruction were embedded with the content of the turn or episode. Teacher 63 talked her students through a visualization of a familiar environment and compared this to the informational text the class was reading in order to help students understand the abrupt end of the Mayan civilization. Teacher 64 referred students to a graphic organizer to help them explain the main idea of the fable they were reading as part of a lesson using test prep materials. In both instances, explanations are explicit and extended pedagogical attention to reading strategies that draw on personal experiences and inferencing information not directly stated in the texts to support students' text comprehension.

Case summary: General patterns of teacher explanations.

Across the sample, teacher explanations tended to occur independently, in the sense that there were no patterns of other codes that precipitated or followed these explanations. There were clearer patterns in how teacher explanations related to instructional materials, instructional talk genres, and language arts topics addressed in the lessons. Informational texts were the most frequent to co-occur with teacher explanations (13% of explanations fell within these episodes). It would seem that instruction using informational texts was frequently supported by explanations. Text comprehension was the most frequent topic in which explanations occurred. Most of these explanations occurred during the Q & A talk genre (43%), with small group text talks the next most frequent genre (16% of explanations occurred during these text talks). These patterns suggest that explanations occurred within many of the most frequent global feature codes (see Appendix B and previous section of this chapter).

Most teachers who used high-rates of explanations used a consistently high number across their three observations, except Teacher 4, who used many explanations in her first observation, and none in her second (Teacher 4 was one of three teachers to only have two observations). The length of explanations in this "high-rate" group of classrooms varied. Teacher 27 used explanations that were usually no more than three lines of transcript in her lessons, where as Teacher 23 had long blocks of uninterrupted talk between questions and brief student contributions. Teacher 10, the only SEI teacher in the sample, was the only teacher in this group to use explanations to build students' background understanding before they read a story together. The teacher prepared students for the text by describing the theme of the text, and by describing content of the main events in the stories.

Teacher explanations tended to fit into the IRE pattern and replace teacher evaluations in the 'E' position when an evaluation was not used. The difference between teacher evaluations and explanations was that explanations were not characterized by reciprocity, since by definition, explanations do not respond to student questions or elicit responses of students. With the good level of coding reliability ($\kappa = 0.69$), teacher explanations were reliably identified talk moves that predicted higher than average reading comprehension outcomes. The content and function of explanations however, disconfirm the study's hypothesis that dialogic talk moves aid students reading comprehension.

Case 2: Uptake questions.

A total of 1541 uptake codes were applied to the transcript data, accounting for 7% of all codes used. Across the sample, the range of uptake questions used by teachers was fairly consistent. Most teachers posed much less than one uptake question per minute of instruction, with an average rate of 0.32 questions per minute ($SD = 0.19$), or about 19 uptake questions per hour of instruction. Teacher 73 was the one outlier that used almost one uptake question per minute of instruction ($M = 0.99$, $SD = 0.42$), or about 59 uptake questions per hour. The next highest rate was Teacher 26 at an average of 0.64 uptake questions per minute ($SD = 0.27$). To get a sense of the range of average uptake rates across the 31 classrooms, a visual of the distribution of average rates of uptake across the classrooms is presented in Figure 5.1.

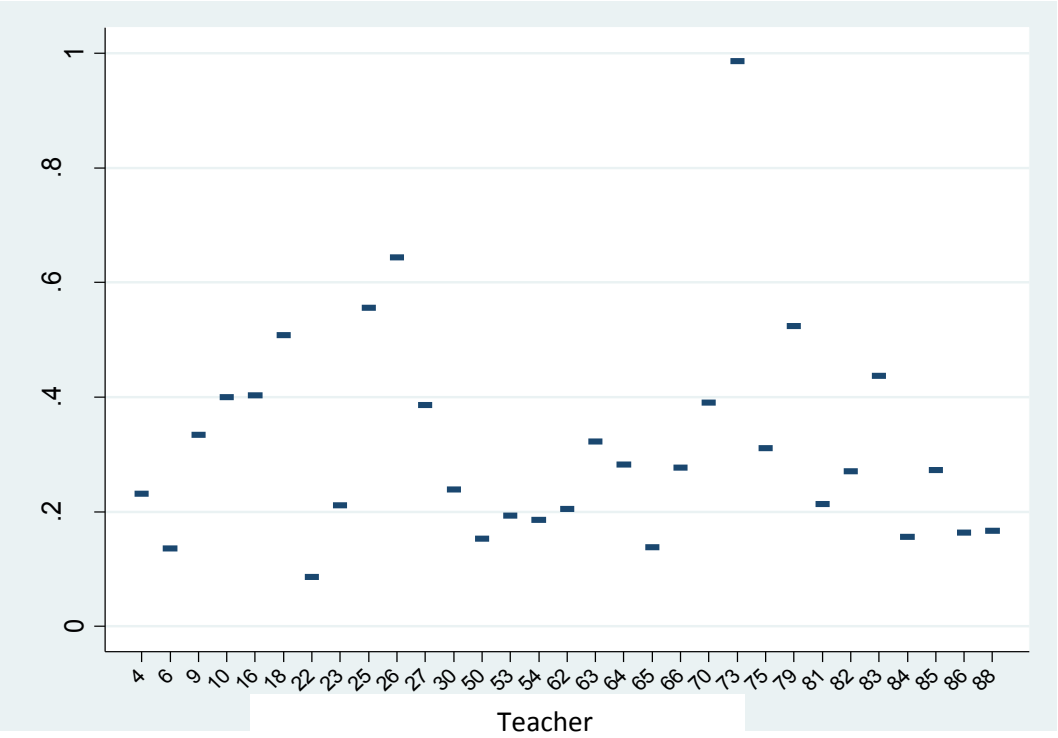


Figure 5.1 *Distribution of Average Rates of Uptake Questions across Teachers*

Fourth-grade teachers showed more variability and a higher median of uptake questions (median = 0.39) than third- and fifth-grade teachers in the sample (medians = 0.21 and 0.25, respectively). Differences between fourth- and fifth-grade teacher questioning were statistically significant: fourth-grade teachers used significantly higher rates of uptake questions ($t(142) = 5.05$, $p < .001$) than fifth-grade teachers across the sample. There were no significant differences in the rate of uptake between third- and fourth-grade classrooms, or between third- and fifth-grade classrooms, and there were no significant differences between the two school districts.

In the literature, uptake questions are defined as questions that focus instructional attention on a previous comment or contribution to press for more information (McElhone, 2012; Nystrand et al., 2003), contingent on preceding utterances (Boyd & Rubin, 2006). In the current study, uptake questions demonstrated a reciprocity in talk between the teacher and one or more

students, and this reciprocity tended to be controlled by the teacher, as it was s/he who chose which student contribution to focus on. In doing so, s/he also let the student(s) know that their contributions warranted further attention. This attention varied in the quality of the uptake question. Multiple reads and coding of transcripts with a focus on this talk move and the surrounding discourse were done, and four themes were refined from the inductive coding. Uptake questions functioned in four ways to encourage and manage this reciprocity of instructional talk. Listed in descending order by the frequency of usage, teachers used uptake questions to 1) press students for explanations of their thinking, 2) encourage students to provide more known information, 3) extend the focus of the lesson by using uptake as a base or foundation to move the lesson forward , and 4) elicit participation from more students. Each theme is discussed in turn below.

Uptake as a press for explanations.

The most frequently used uptake question picked up a student's contribution, usually in answer to a previous teacher question, and extended this by using questions like 'what' and 'why'. These questions pushed students to explain their thinking behind their previous contribution in more detail. Table 5.8 presents some typical examples of uptake questions that pressed students to explain their thinking.

Table 5.8 *Typical Examples of Uptake Questions that Press for Explanations*

Teacher	Speaker	Discourse and Uptake *(in bold)*
81	Student:	He was a great athlete.
	Teacher:	**What do you mean by great athlete? What made him great?** You're right, he is a great athlete, but what made him great? You know this Iman, you talked to me about him all last week. You know this. You can do this. What did you learn about him? (P108:207, 3/3/11)
83	Teacher:	Okay, what else did we say about Mr. Yeager? What other attribute or characteristic did we give him?

	Student:	Intelligent.
	Teacher:	He was intelligent. **Why did you say that, that he was intelligent?**
	Student:	Because he knew that he was hurt so he just took like a broomstick so he could flip the switch because he knew that his ribs were hurting and he just lean forward.
	Teacher:	Good. He broke his ribs. Did he show people, oh, I have a broomstick and I'm gonna break—no, what did he do? (P32:9, 1/2/11)
25	Student:	What in the world is a tall tale?
	Teacher:	**Why do you think they call it a tall tale? Kevin just said 'what in the world is a tall tale?'** [*Inaudible student responses*]
	Teacher:	**Why do you think [that], James?** Think of the things that happen in the *Heat Wave*.
	Student:	Goes over the edge?
	Teacher:	Yeah, it goes over the edge. It makes things really, really exaggerated, like [*references story*] the geese flying through the clouds and coming out stuffed, plucked, and roasted. (P41:217, 2/10/11)

Note. Words in bold represent coded instances of uptake questions

In each of these examples, the teacher used the students' words almost verbatim to affirm the students' answers and then extended them into follow-up questions. Teachers 81 and 83 both used an uptake question to press their students to explain their previous answers about the texts they were reading, encouraging their students to justify their contributions by citing more facts from the texts. These questions indicated a pedagogical attention to student text comprehension by soliciting more depth to their previous answers. Teacher 83, reviewing previous work with her students, sought pre-specified answers with her test question (e.g. What other attribute or characteristic did we give him?), but within this Q & A pattern, which used many test questions like this, the teacher managed a brief reciprocity of ideas with her follow-up question, pressing the student to explain his evidence for his answer. Rather than answering a student's question directly, Teacher 25 turned a student's question back to the class to answer, eliciting participation (Theme 4) and then pressed the student to explain his thinking. Teacher 25 then extended the student's answer by providing the synonym 'exaggerated' and some textual

evidence. In three turns, Teacher 25 not only affirmed the student's question, but elicited more participation by the class as a whole, elicited more talk from one student in particular by pressing him to explain his thinking, and then recasting the student's answer with a synonym and provided an example of the synonym from the story.

Despite the emphasis of these questions on eliciting explanations, very few of uptake questions elicited extended student talk, which, in this study, was coded as "student explanations" spanning 2 or more lines of transcript or addressed 2 or more ideas (see Case 4 in this chapter). Using search queries in Atlas ti 6.2 software, only two percent of uptake codes were followed by student explanations. One reason for this may be in the kinds of explanations the teachers frequently sought. Each uptake question was coded for the quality of thinking it elicited from students, coded as either "high-cognitive quality" or "low quality" questions (see Appendix A coding scheme). High-quality uptake questions were those that asked student to infer/speculate, analyze, evaluate, or generalize information. Low-quality uptake questions were questions that had students recite or report information that was presumed knowledge. There were 498 uptake questions that were coded as high-quality questions out of a total of 1541 uptake questions (32%). The rest (1043) were coded as low-quality uptake questions (68%). Both kinds of uptake questions occurred in all lessons. It is noteworthy that uptake questions and high- and low-quality questions were strongly correlated with each other ($r = .80$, $p < .001$ and $r = .75$, $p < .001$, respectively). These relationships will be discussed in Chapter 6.

High-quality uptake questions tended to ask students to infer or speculate on important ideas from stories and to make predictions about plot lines. There were very few instances of uptake questions that had students analyze or evaluate information about texts or language arts

content, which may have elicited more extended turns of student talk. Some typical examples of high-quality uptake questions are included in Table 5.9.

Table 5.9 *Examples of High-quality Uptake Questions*

Teacher	Speaker	Discourse and Uptake *(in bold)*
27	Teacher:	What does that mean? Put it in your language. It was more of a task than she had supposed. Mike.
	Mike:	It was more of something to do than she thought it was.
	Teacher:	Very good. It was more of something to do than she thought it was. **Does that show that she's adventurous?** Thumbs up, thumbs down. If you're not sure, put [*unintelligible*]. She could have said this is way more than I expected, no way am I going to try and put together this clock. I think that's pretty good evidence.
	Student:	[*unintelligible*] willing to take risks.
	Teacher:	They are willing to take risks. **Is there enough evidence showing that she's willing to take a risk?** [P1:81, 1/28/11]
73	Student:	Oh!
	Teacher:	What? What's oh? **Tell me what you're oh-ing.**
	Student:	He's mentioning all the colors that you can have in your—
	Teacher:	**Why would he do that?** Did you guys notice that?
	Student:	[*Inaudible*]
	Teacher:	**What does that mean?**
	Student:	It means—
	Teacher:	Why would it be colors?
	Student:	--people can be any color.
	Teacher:	**What? Say that again?** [P20:106, 3/3/11]
81	Teacher:	The main word inside the whole word. Okay. Not a prefix, not a suffix, the main word. So, in the word, "disappointment," what is the main word in that one? Yes?
	Student:	Appointment.
	Teacher:	Okay. You're saying, "appointment." **Why do you say appointment?**
	Student:	Because the root or base word, most of the time ends in an N.
	Teacher:	So you knew that "dis" was what?
	Student:	A prefix.
	Teacher:	A prefix. Dis is a prefix. Do you have a suffix in this word?
	Student:	Yes.
	Teacher:	Yes. What's the suffix?
	Student:	Ment.
	Teacher:	Ment, that's the last part of it. So what's my base or root word?
	Student	Appoint [P106:16, 2/3/11]

Note. Words in bold represent coded instances of uptake questions

Teacher 27 was intent on having her students back up their inferences with textual evidence. As the class constructed a reading response paragraph together, the teacher took up two of her students' contributions to make sure students agreed that there was enough textual evidence to support both inferences about the character. Teacher 73 asked her students to infer the meaning of a short poem, and to repeat a response for the whole class to hear. However, student contributions were interrupted twice, attenuating any elaborated student talk. Teacher 81 used one uptake question to solicit a student's explanation of the root of the word 'disappointment.' When the student replied with a mistake in her reasoning, the teacher guided her to the correct answer by posing test questions in the form of fill-in-the-blank ("So you knew that "dis" was what?") and yes/no questions ("Do you have a suffix in this word?"). In this example, the uptake question elicited only a sentence of explanation from the student, likely because of the content, and then Teacher 81 heavily scaffolded the student to reconsider her initial answer without explicitly evaluating the answer as incorrect.

There was some evidence that uptake questions during writing conferences (a rare activity in these observations, and usually done one-on-one or in small groups) may have elicited some extended student talk in the form of writing. In the follow two exchanges, teachers tended to read student work and take up the student's written ideas orally to press the student for more information or scaffold student ideas. The student with Teacher 23 does this orally, while the students of Teacher 66 presumably explain their ideas more fully in their written text.

> Teacher 23: So it's very interesting for me to hear [read] about this boys' party. **What makes it such a boys' party?**
> Student: I told my mom that I want a boy's only party and so I just invited all the boys to the party.
> Teacher: What about your sister?
> Student: She wasn't born yet (P76:44, 2/9/11).

Teacher 66: I also want you two to think about, Natalia and Michelle 27:16, why was she untruthful and why—how did we know that she was worried about [inaudible 27:25]? Was she untruthful because she was sneaky and a bad person? Why was she being untruthful?
Student: She said that she had no problem [inaudible 27:35] she was terrified of him.
Teacher: **Why? Why did she do that?** (P68:80, 1/12/11).

In conclusion, uptake questions as presses for explanations indicates a pattern of reciprocal discourse that was heavily guided by each teacher, so it would seem that this give and take is more constrained than a typical discussion by two adults, for instance. Despite a lack of extended student talk following these questions, teachers did work to centralize student ideas, which were not only a pedagogical focus during that moment of instruction, but also were opportunities for teachers to press students about their thinking or hold them accountable for further explanations, which provided some opportunities for extended talk and analytic skills like justifying reasons and providing evidence from text, a key skill on the reading comprehension measure used in this study.

Uptake as a press for more information.

A second theme identified in the analysis of uptake questions was the use of these questions to encourage students to provide more known information. The majority of these uptake questions occurred during instruction around text comprehension, and less frequently, during form-focused lessons on grammar worksheets. Table 5.10 presents some typical examples of uptake that elicited recalled or known information.

Table 5.10 *Typical Examples of Uptake Questions that Press for Known Information*

Teacher	**Speaker**	**Discourse & Uptake** (in bold)

10	Teacher:	**What word are you looking for? Tell me your idea, what do you think is going to happen?**
	Student:	The little girl goes to school.
	Teacher:	Oh, the little girl might go to school. Well, he's in Japan and the girl is old now.
	Student:	Married?
	Teacher:	**Who's going to get married?**
	Student:	The daughter?
	Teacher:	**How do you know that?**
	Student:	The picture
	Teacher:	The picture. Ok, let's listen and find out. (P4:69-70, 12/20/10)
27	Student:	I think it's supposed to be kind of like sum all the evidence and details and then [noise obscures] like an ending.
	Teacher:	OK. So part of it, yes, **it's summing it up. What's another way to say that?**
	Student:	[*Unintelligible*]
	Teacher:	Your closing sentence, your conclusion, how are you going to conclude. Yes. (P1:12, 1/28/11)
25	Teacher:	Yep and what do you do when you go back to the story?
	Student:	Underline.
	Teacher:	**Underline the what?**
	Student:	The information.
	Teacher:	Okay, or the --? **What's that information?** What is Andrea talking about?
	Student	The answer. (P42:147, 2/15/11)

Note. Words in bold represent coded instances of uptake questions

In each example in Table 5.10, teachers took up student contributions to elicit information that was presumed knowledge. Teacher 10 taught a Sheltered English Instruction (SEI) fourth-grade class of English learners who were mainly Spanish speakers (one Portuguese speaker is noted in the transcript). Her uptake questions in each of her three observed lessons were consistently used to clarify and assess students' understanding of a text they were reading together. Uptake questions were focused on vocabulary and plot comprehension, picking up a student's answer and scaffolding toward a more complete and correct answer. These took place in a whole-class, Q & A genre format that used a basal reader and a vocabulary graphic organizer. In the above exchange, uptake questions acted to heavily scaffold the exchange between the teacher and one student, and, it seemed, to check for student comprehension

("Who's going to get married?") and evidentiary support ("How do you know that?"). The student's verbal responses were quite attenuated, due in part by the closed question ("Who's going to get married?") but may also have been due to the student's English speaking proficiency.

Teacher 27 used uptake in a similarly constrained fashion by seeking a vocabulary term that was assumed knowledge ("What's another way to say that?"), as indicated by her evaluation of 'yes' when a student provided the term she was looking for. Teacher 27 accepted her student's answer and took it up as a basis for her next question, a brief 'spot check' of students' academic vocabulary. Teacher 25 was observed to use form-focused worksheets in all of her observations, and the questions were similar in quality. Students' answers were taken up to extend the correct answer into another known answer, such as asking a further question about the grammatical concept or retelling a plot line with more accuracy than originally offered.

Uptake questions that acted to encourage more information from students, like those that pressed students for more explanations (Theme 1), did not generate elaborated student talk. In contrast to Theme 1 uptake questions, however, this pattern was predictable since most of these solicitations for more information were coded as low-quality questions that required students to recall assumed knowledge. Low-quality uptake questions tended to be factual questions about text comprehension during text talks that did not require student explanations, analysis, or inferences. The following examples (Table 5.11) clearly show the attenuated nature of student talk when uptake questions for information were posed, while still exemplifying the reciprocal nature of these talk moves.

Table 5.11 *Examples of Low-quality Uptake Questions*

Teacher	**Speaker**	**Discourse and Uptake** *(in bold)*
26	Teacher:	So what do they become when they become naturalized?

	Student:	A citizen.
	Teacher:	**A citizen of?**
	Student:	Of the United States. [P51:25, 12/14/10]
50	Teacher:	What next? He writes silly stories; he always uses his good luck pen. What happens next? Aaron?
	Student:	He lost it.
	Teacher:	He loses or he lost—**what's he lose? Loses it like goes crazy, loses it?**
	Student:	No. He lost his good luck pen. [9:114, 3/3/11]
63	Teacher:	What are you writing? **Polar bears and arctic what?**
	Student:	Foxes.
	Teacher:	**What do they do?**
	Student:	I don't remember
	Teacher:	You don't remember? …You highlighted it. What do polar bears and arctic foxes have? [24:90, 1/13/11]
64	Teacher:	Where does the story take place? Are we on the moon? Are we in the mall? Where are we, Katie? Where is the story taking place?
	Student:	Outside.
	Teacher:	It is outside. **Can you give me a little bit more information, [another] Child's Name?**
	Student:	It's at a pond or a river.
	Student:	Or a shack.
	Student:	A pond or a river?
	Teacher:	A pond or a river, yes. So the setting we can say is –is a river… [34:35, 1/14/11]

Note. Words in bold represent coded instances of uptake questions

In all four examples presented in Table 5.11, uptake questions ask each student to provide more information about the text under discussion, factual information that required the student to retell what s/he had just read or highlighted. The uptake question from Teacher 26 was a simple fill-in-the-blank question, and Teacher 50 asked a clarifying, either/or question of her student to better identify which meaning of the expression 'lose it' he is using. Both were seeking specific vocabulary in order to check students' comprehension of the texts, and thus required little language other than one word. The other two examples of uptake were not as structured, allowing students the opportunity to respond with more than one specific term, but both teachers'

evaluations of the students' responses suggest these uptake questions were asking students to retell ("What do arctic foxes do?") or recite known information. For instance, Teacher 64 sought 'river' as a more specific answer than 'outside' to her uptake question.

Like uptake as a press for explanations, these uptake questions also occurred in a pattern of reciprocal discourse that was heavily guided by each teacher, and indicated a lack of a pattern eliciting extended student talk. Despite a lack of extended student talk following these questions, teachers continued to centralize student ideas, but rather than presenting students with opportunities for articulating their thinking, these uptake questions provided the teacher, and presumably the class, with correct answers with more specificity than the original contribution. Uptake questions that pressed students for known information emphasized accuracy over explanations, analysis, or synthesis of information.

Uptake as extension.

Another theme from analyses of uptake showed teachers taking up students' contributions and using them as foundations or starting points to move the lessons along. Similar to the previous theme of uptake questions that elicited more information from students, these segments of talk seemed to also solicit information, but to be used to extend the lesson in ways firmly managed by the teacher, rather than simply have the student provide a more accurate answer. Most cases of uptake that extended the lessons were used with unabridged and basal narrative texts, and were used to take a student's answer to further along the talk about the story. For instance, Teachers 50 and 22, both third-grade teachers, posed the following questions to assess how their students understand the text: "Can someone tell me more about (what the character does that was just provided by a student)?" (P9:184, 3/3/11), and "Rashid, what would you say then? What else would you add? (P17:24, 3/2/11). These differed slightly from asking

students to further explain their contributions according to the content to which the teacher's sought (Themes 1 or 2): "Why though? I want you to tell me why?" (P17:16, 3/2/11). In this latter example, the momentum of the lesson slowed to accommodate a student's explanation and justification of her answer, in contrast to the former examples which clearly intended to elicit more answers, rather than more details to answers, to move the talk about text forward.

For instance, Teacher 18 did this repeatedly when conducting a whole-class discussion on the main character of *Esperanza Rising* (Ryan, 2000), a novel about a young girl coping with a loss of social status when immigrating to the US from Mexico (Table 5.12).

Table 5.12 *One Teacher's Use of Uptake Questions as Extensions Guiding a Whole-class Text Talk*

Turn	Speaker & Discourse (*uptake in bold*)
1	Teacher 18: So Sahira just said, if you couldn't hear her, that in Mexico, she [the character] had to do nothing. She didn't have to work. She got to, you know, lazily walk around the ranch, look at the roses, right, eat the grapes. Now, what does she have to do? **Does she have that lifestyle?**
2	Student: No.
3	Student: No.
4	Teacher: She's got to pitch in, right? She's got to pay her share. She has to do everything herself. Do you think that was hard for her?
5	Student: Yeah
6	Teacher: How did that change—oh, she definitely didn't like that. How did that change her, do you think? Do you think anyone laughed at her in Mexico?
7	Student: No.
8	Teacher: **Why not?**
9	Student: Because she was [*inaudible*].
10	Teacher: **So how do you think she felt** [*inaudible*]?
11	Student: She would feel like she [*inaudible*].
12	Teacher: How did the poor girls feel about Esperanza in Mexico?
13	[*Cross talk*]
14	Teacher: **All right, so what is that? Is there a word for that?** You gonna [*inaudible*] for what someone has? [*Cross talk*] Yeah, some of the girls were jealous of her in Mexico, right? Envious of her. How do they feel about her now?
15	Student: Nice to her.
16	Teacher: **Are they nice to her, now?**
17	Student: No.

18	Teacher: They're laughing at her, watching her try to hold the broom. [P79: 2/26/11]	

Note. Words in bold represent coded instances of uptake questions

In this segment of talk, each uptake question indicated a reciprocity where the students gave and the teacher took ideas to move the lesson along, a lesson focused on seeking evidence of student comprehension of a main theme of the novel. In the first turn, Teacher 18 focused the class' attention on one student's contribution, which led to the teacher posing questions to elicit comparisons of the character's life before and after her immigration to the US. Her question pressing a student for an explanation (line 8) generated only a line of student talk (that was unfortunately unclear in the audio), and followed this 'press' with two more (lines 10, 14). These uptake questions culminated in enough information for Teacher 18 to end her questions eliciting comparisons of the character's two lives, and to cite evidence for the student's answer (line 17).

It is interesting to note in this segment of instructional talk that the teacher did not use low-quality evaluations of student answers (e.g. "Yes, that's correct"). Teacher 18 used questions as suggestions for answers (line 4) and uptake instead of evaluative comments to further the talk gauging students' comprehension of the theme. This was in contrast to uptake that pressed students for more information (Theme 2): encouraging students for more information tended to elicit known facts and were followed by low-quality teacher evaluations.

A final characteristic of uptake questions that extend from student contributions to move the lesson forward was the frequency that these occurred with instructional episodes coded as read alouds. This pedagogical talk genre, of reading (narrative) texts aloud as a class, seemed to frame uptake questions by the need to keep the story moving along. In these lessons, uptakes functioned to build understanding in a linear fashion, rather than to elicit more depth of thinking or talking. This lateral movement of collective instructional talk allowed for the reading of the

text to clip along at a consistent pace, punctuated by teacher questions attending to comprehension of the plot and character motives. Two examples are presented in Table 5.13.

Table 5.13 *Examples of Uptake Questions as Extensions during Read Aloud Episodes*

Teacher	Speaker	**Speaker & Discourse** (uptake in bold)
53	Teacher:	…Touch therapists can detect the energy field and heal people. Does Emily believe in touch therapy or is she skeptical of touch therapy?
	Student:	She didn't know—
	Teacher:	Tyrone?
	Student:	She believed in it.
	Teacher:	**Did she believe in it? Let's go back and look. She's saying that she wanted to test whether it was true or not. So does she believe in it yet?**
	Student:	No.
	Teacher:	No, but she's going to find out. Right, she's doubtful. Remember our word yesterday. We talked about skeptical meaning you're doubtful. All right, let's look on page 216. Flip the page. Who would like to read the first paragraph, 216. (P47:156, 2/1/11)
50	Teacher:	…He's sad and he thinks he can't write. He's not a good writer; it's only because of his good luck pen. Then what? Curtis, then what happens when he's so sad, he thinks he can't write. What's next?
	Student:	He gets another pen.
	Teacher:	**How does he get the other pen?**
	Student:	From his friend.
	Teacher:	Oh, so Tricia gives him what? What, Curtis, gives him what?
	Student:	A pen.
	Teacher:	**A new pen. Does she just give it to him and say, "Here you go"? Or does she just give it to him and say, "Eh, you're a bad writer**"? Does she just give it to him and say, "Happy birthday"? She says something pretty important. I know one, two, three people—four people have read
	Student:	that story. Omar, what happens?
	Teacher:	[*Inaudible*]. **How does he learn that? Does she say anything to him?** (P9:120, 3/3/11)

Note. Words in bold represent coded instances of uptake questions

Both teachers in Table 5.13 take up student answers to verify their comprehension of the stories' plotlines or characters' motives, doing this in such a way that the read aloud continues the linear progression of the read aloud with little break between questions. The focus of the lesson segment from Teacher 53 seems to be more on questioning than the content of the

students' thinking underlying his (incorrect) answer. The linearity of instructional talk to move through the text was also a clear pattern in the pedagogical discourse of form-focused worksheets. Form-focused worksheets were more frequently used in classes that had low rates of uptake questions (in the 25th quartile), and these questions tended to be of low-quality, eliciting known information from students. This kind of highly-structured text may limit reciprocity, or may shift classroom discourse to a focus on eliciting answers that are specific and can move the lesson through the worksheets expeditiously. Teacher 6 provides a typical example of this pattern of talk structured by worksheets.

> Student: [*reading her answer on worksheet*] I played baseball an hour ago.
> Teacher 6: **I played baseball an hour ago. What's the verb? What's the subject?** Jacob?
> Student: Baseball. No. [*unintelligible*]
> Teacher: How about let's shorten it. I play baseball. What's the subject of the sentence? Tanisha?
> Student: Play?
> Teacher: **That would be the--?**
> Student: Verb.
> Teacher: Verb, because it's an action. You're playing. But what's the sentence about? (P111:113, 1/28/11)

Again, like the other uptake themes, uptake questions as extensions occurred in a pattern of reciprocal discourse that was teacher-managed and showed a lack of extended student talk. The main feature of these uptake questions was that they functioned to incorporate student contributions in the interest of forwarding reading and form-focused lessons. Teachers centralized student ideas to build upon them to move to the next question or plot event, rather than explore the student's thinking. Like uptake that elicited known information, this form of uptake pressed students for accurate information on which to proceed with the teacher's next point.

Uptake as elicitation of participation.

The fourth and final category of uptake questions identified in this analysis was the most infrequently used. A few teachers took up student contributions to focus the class and bring multiple students in to consider the question, rather than press a single student to explain or extend his or her answer. Three examples of uptake as participation are listed in Table 5.14.

Table 5.14 *Typical Examples of Uptake Questions that Elicit Participation*

Teacher	Speaker	**Speaker & Discourse** (uptake in bold)
54	Teacher:	…What made them vibrant? Nicholas?
	Student:	They have bright colors?
	Teacher:	They have bright colors. **Please stand up if you are wearing even the littlest bit of a vibrant color. Who thinks they have a little bit of vibrant color on?** (P63:122, 1/31/11)
27	Teacher:	Very good. It was more of something to do than she thought it was. **Does that show that she's adventurous?** Thumbs up, thumbs down. If you're not sure, put [*unintelligible*]. She could have said this is way more than I expected, no way am I going to try and put together this clock. I think that's pretty good evidence.
25	Teacher:	**Why do you think they call it a tall tale? Kevin just said 'what in the world is a tall tale?** [*Inaudible student responses*]

Note. Words in bold represent coded instances of uptake questions

Each of these examples shows an intention to garner more student participation in the content of the talk. Teacher 54 took up a student's answer by having the class get somewhat kinesthetically involved, and she was able to personalize the vocabulary lesson in three turns of talk. In addition, Teacher 54 was one of the few teachers that used uptake questions to draw on students' personal experience related to the vocabulary under study. Teacher 27 takes up a student answer to briefly check students' thinking without eliciting any student talk, and Teacher 25 takes up a student's question to pose it to the entire class. This is perhaps the most clearly articulated example of the reciprocity embedded in these kinds of questions. Teacher 25 diverges

from her extremely consistent pattern of teacher-led Q & A across all three of her lessons to share a student's pertinent question with the whole class.

Without more examples of uptake as participation, it is difficult to determine if this form of uptake can elicit in-depth thinking or higher-order thinking skills, or whether it may serve to maintain the lesson momentum. However, it does seem like this form shares one similarity with the other forms of uptake presented above: they function to focus students' attention on either their peers' talk about text or the text itself. This management of student attention is discussed more in Chapter 6.

Case summary: General patterns of uptake questions

Uptake questions were significant predictors of student reading comprehension achievement, confirming the hypothesis that this dialogic feature of instruction is positively related to reading outcomes. Teachers posed these questions to press students for more explanation of their thinking, to encourage students to provide more known information, to push the lesson forward from a 'base' of a student's contribution, and to seek more participation by turning a student's question or contribution over to the class (though the latter was rare). Whether the uptake questions were of high or low cognitive quality, they were predominantly focused on exploring and verifying students' text comprehension. Across all 88 transcripts, there were no clear patterns of uptake questions preceding or following certain talk moves. In three percent of the instances of uptake, these questions followed low-quality teacher evaluations or test questions. In addition, uptake questions did not tend to cluster together. However, there were patterns at the instructional episode level of coding.

Uptake questions occurred in teacher-managed instructional episodes that predominantly focused on text comprehension of narrative texts (unabridged and basals; 49% of all instances of

uptake codes) and testing materials (22% of uptake codes occurred with the testing material code). When working with form-focused and vocabulary worksheets, teachers rarely posed uptake questions (5% and 2%, respectively), which was the main difference in coding patterns between high uptake-rate and low-rate classrooms. Whether the uptake was a press for an explanation or a press for more known information, the vast majority of these questions focused students' attention on main ideas and details of the plot and characters' motives in narrative texts and test passages. These patterns suggest that an instructional focus on the comprehension of narrative texts and test-prep materials may lend more opportunities for some reciprocity between teachers and students, but it is also important to note that more uptake questions happened during instances of the two most-used materials across the entire sample. Similarly, the Q & A talk genre was by far the most used instructional talk pattern seen across the classroom observations, including high- and low-rate classrooms, and 45% of uptake questions occurred in episodes coded with this talk genre. Read alouds and small group text talks also showed high frequencies of uptake question use.

Despite the frequency of uptake questions as press for student explanations of their thinking, very few of the student responses in either the mainstream or SEI classes where coded as extended talk. Students across these classrooms did not answer many questions with more than one word or one phrase responses, which may reflect the preponderance of the Q & A talk genre that framed learning activities. This pattern also mirrors the expectations of the reading comprehension outcome measure: students were required to read a sentence or short passage and supply a vocabulary word consistent with the passage topic, drawing on students' expressive vocabulary knowledge and inferencing skill to determine an appropriate word.

Finally, uptake question use was quite consistent within classrooms. Teachers in high-rate classrooms tended to use high rates of uptake across all of their observations. The main difference between the two groups was the use of worksheets: very few teachers who used worksheets used many uptake questions.

Case 3: Low-quality teacher evaluations.

Low-quality teacher evaluation moves were positively associated with students' reading comprehension scores. They accounted for 12% of the codes applied to the 88 transcripts in the sample. The average amount of low-quality teacher evaluations (TEvalLs), without accounting for the length of the observations, was 30.53 evaluations per lesson, with a large standard deviation of 20.00. Every classroom used at least one TEvalL. The average rate across the 31 classrooms was 0.54 (*SD* = 0.22) or 32 evaluations per hour of instruction. Fourth-grade classrooms showed the highest mean rate of this talk move at 0.64 TEvalLs per minute of instruction. Teachers in the Massachusetts sample showed a higher rate of these talk moves, with an average rate of 0.62 (*SD* = 0.26) low evaluations per minute, and a higher median of 0.60 TEvalLs per minute. Teachers in the Maryland had a lower average rate (*M* = 0.50, *SD* = 0.20) and a lower median score of 0.44 low-quality evaluations per minute. These talk moves were the most widely applied code to these transcript data.

TEvalLs in this study were defined as teacher comments about student contributions that repeated a student's response verbatim, expressed praise ("Good. That's right"), or sympathy ("I wouldn't like that either"), coded teacher turns that indicated no audible evaluation of student's response. In these data, this was infrequent. The qualitative differences between TEvalLs seemed to be with what proceeded from the evaluation. Some TEvalL moves were simply affirmations and the lesson moved on to a new focus with the next teacher question. Other evaluations

conditionally accepted answers and moved the lesson along, in many cases asking other students to contribute so that a more acceptable answer became apparent during the turn taking. A third characteristic of TEvalLs was their function as a short, precipitating talk move before building onto the topic, usually with a question related to the "affirmed" or accepted student contribution. This was the least observed of these characteristics of this talk move. Thus, across the classrooms and grade levels, TEvalLs reflected three themes capturing their content and function during instructional talk, listed in descending order of frequency: TEvalLs 1) as affirmations, 2) as building blocks for extending talk, and 3) as conditional affirmations.

Low-quality teacher evaluations as affirmations.

The majority of TEvalLs in this sample functioned to affirm students' contributions. Most TEvalLs functioned to accept students' contributions into the lesson discourse, mainly by affirming acceptable answers and continuing with the lesson discourse genre. These evaluations did not function to "reformulate" students' responses to scaffold content (McElhone, 2012), and did not reconceptualize contributions to provide students with an opportunity to think about the topic in a new way or connect it with previous understandings (Cazden, 1988). Instead, these evaluations acted as perfunctory teacher acceptance of the student's thinking before moving the lesson along. They worked mainly to let students know that their brief answers within the pattern of talk constituted correct content, and that the pattern of talk could move forward. In this sense, when following how teachers addressed text comprehension (the most frequent topic of instructional episodes in which TEvalLs occurred), these evaluations allowed, for a brief moment, students' contribution to be the content of the instructional talk, and signaled that this content was acceptable to the topic being addressed. Typical examples of this talk move are presented in Table 5.15.

Table 5.15 *Typical Examples of Low-quality Teacher Evaluations as Affirmations*

Teacher	Speaker	**Discourse and Low-Quality Evaluations** (in bold)
9	Teacher:	…Okay, who else has to share? Floria?
	Student:	[*reading her writing*] Tears were streaming down my face and my frown was beginning to be worse.
	Teacher:	**Her frown was getting worse.** Yes?
	Student:	Sad.
	Teacher:	**Sad**. Yes?
	Student:	Depressed.
	Teacher:	**Depressed. Could be depressed**. Is there someone that I didn't get to yet? Go ahead and then I wanna [*cross talk*].
	Student:	I have two.
	Teacher:	Go ahead. …
	Student:	My face was as red as a plumped out strawberry.
	Teacher:	**Face was as red as a plumped out strawberry.**
	Student:	I was like a volcano ready to explode.
	Teacher:	**A volcano ready to explode, oh no!** Devon?
	Student:	Nervous?
	Teacher:	**Well, I'm thinking something different.** Floria?
	Student:	Angry. (P85:88, 1/25/11)
79	Teacher:	There's a lot more text pieces you have to pay attention to in this story.
	Student:	One page looks like it's kinda broke up in how many parts?
	Teacher:	Three. **Three parts there.** What about over here; how many parts are they broken up into? (P103:11, 2/2/11).
6	Teacher:	[*classroom management talk*] … Let's just really quickly go through these together to make sure we're all on the same page with this. Okay. Number one, 'let us.' What would that be?
	Student:	Let's.
	Teacher:	**Let's.**
	Student:	Can we correct this with a pen? Can we correct this with a pen?
	Teacher:	You can just fill it in with your pencil if didn't have it. (P113:61, 2/18/11)
54	Teacher:	You retreat, you go back. Not back in time, not time travel, but you are leaving a situation that you don't think that you are going to be successful or win. Yes, Cam?
	Student:	Like fleeing?
	Teacher:	**Like fleeing. Excellent synonym, Colin. Excellent.** (P63:52, 1/31/11)
25	Teacher:	Anything you can tell me about poetry. [Put it away.] What do you know about poetry? Yes.
	Student:	It sometimes rhymes.

Teacher:	**It sometimes rhymes,** Kendall?
Student:	It doesn't always rhyme.
Teacher:	**It doesn't have to rhyme, it doesn't always rhyme.** What else do you know? (P42:38, 2/14/11)

Note. Words in bold represent coded instances of low-quality teacher evaluations

Teacher 9 had the students focus on their writing of details that show a character's feeling without explicitly identifying the feeling (eg. Ramona felt "like a volcano ready to explode" because her sister ruined her painting). In this excerpt, students shared their sentences and their peers were asked to infer the emotion. TEvalLs punctuated each of the teacher's turns, used to affirm students' contributions as the students took turns sharing their sentences. Most of her TEvalL were simple repetitions of students' contributions. One affirmation suggested some uncertainty ("Could be depressed"), but this was not picked up by the teacher. In the teacher's last turn, it became evident that she was seeking a specific adjective ("Well, I'm thinking something different"), which turns out to be 'angry'. The evaluation did work to forward the lesson by reformulating the talk pattern from contributions to guessing the teacher's idea, but did not contribute any content or scaffolding of content or language, and thus was coded as another TEvalL.

The short turn by Teacher 79 is a typical example of how TEvalLs precede questions, acting as affirmations of the response, but continuing to move the lesson along. There was no building from the student's response, but a continuation of the Q & A genre of talk for this instructional episode. It is clear with the examples from Teachers 79 and 6 that the lesson format structured these sorts of evaluations. Both teachers seemed intent to cover the material with a Q & A pattern of talk that adhered quite strictly to an IRE pattern. The contributions were brief, and the evaluations were similarly brief, and the questioning moved the lesson forward. Teachers

6 and 54 provide typical examples of how TEvalLs function as a simple affirmation through repetition, and then the response is put aside to make way for other issues, like procedural issues in the case of Teacher 6.

Teacher 25 showed very consistent use of TEvalL across all three observations despite differences in pedagogical materials and topics addressed in each episode. A typical turn began with her question, a short student response, and her brief acceptance of the answer before moving to another question; in other words, the classic IRE pattern of pedagogical talk. This pattern of rapid shifts happened when conducting a whole-class text talk on short story, a review of form-focused worksheets, filling out a graphic organizer, or reading test materials. The pattern was striking in its regularity and total amount of evaluations. The excerpt on Table 5.15 began an instructional episode focused on students' comparison and understanding of two poems. The elicitation of common characteristics of poetry was treated as a review and quickly led to a read aloud and then Q & A on the first poem. The questions followed a worksheet very similar in format to the state reading comprehension test, which attended to recall and inference type questions, and were evaluated as correct or incorrect by the teacher. The content of the Q & A stayed close to the text reflecting an efferent stance toward the two poems, reading them for information to answer the questions. Teacher 25's evaluations reflected this by affirming correct answers through repetitions and moving on.

In the excerpts on Table 5.15, no elaborated student talk is indicated in these segments of instructional discourse. Low-quality evaluations in the current study were hypothesized to not indicate or encourage reciprocal and dialogic interactions during lessons because they are defined by their lack of scaffolding content or language, such as revoicing or reconceptualizing student responses. Though these low-level evaluations do not seem to engage students in

dialogue, there is some suggestion that they may be indicative of more student discussion (Nystrand et al., 2003) when teachers use this low-quality evaluations as conversational placeholders to retain some management of the talk, while at the same time, allowing the teacher to get out of the way of student talk. To test this theory, proximity searches were run on the transcript data to explore if TEvalL codes and student explanations or questions occurred as precipitating or proximal events, and if they occurred frequently together across the sample. Of the 1574 instances of TEvalLs applied to the transcript data, 147 co-occurred with student explanations, extended talk turns by students (of which there were a total of 844), meaning that about 9% of TEvalLs were followed by student explanations. Teacher 27 had many of these during her text talks with her class about the main theme of a reading orcharacters' motivations. The following excerpt of a typical text talk of Teacher 27, exemplifying the use of TEvalLs to manage the pattern of talk but allowed students 'room' to contribute extended thoughts.

> Teacher 27: Yes, Marius?
> Student: It's kinda like he has two consciences and his subconscious is going into a day nightmare in a sense, but his normal superego is just trying to keep that away and stuff.
> Teacher: **Right, so yeah, it's his conscious and his subconscious going back and forth.**
> Students: [*Cross talk*]
> Teacher: **Yeah, maybe have that image. Back and forth, back and forth.** Alex?
> Student: When you said that he was focusing on the rocks in front of him, it was like the funniest thing just popped into my mind. And so we were at Maine. We rented a house and it was night, and we were playing Manhunt. There was a tag [*inaudible*]. And so there was a little cliff. It was only this high…. (P3:117, 3/3/11)

Teacher 27 uses repetitions of a student's contribution as simple affirmations, and these two moves are surrounded by extended student talk, one student theorizing what the character is going through, and the other sharing a similar experience to the text. More typically, TEvalLs as affirmations punctuated turns of student talk that were more brief that the previous example.

Teacher 83: Good, so they started their competition. How was it going? What is happening with these students that had started a competition?
Student: When they were, they usually be louder and louder. Yeah, louder but they usually don't talk. But then one day I watched [*inaudible*].
Teacher: They weren't loud. I think just a couple of the boys and girls spoke out. But what was recess like?
Student: It was like all of them were loud. They were just making noises. They said that they couldn't talk.
Teacher: **Very good. Mm hm.**
Student: Making noises. But they started making noises.
Teacher: So good to the teachers that were watching them, they heard just noises. But what were they? (P29:70, 12/10/10)

Teacher 83 uses a simple affirmation ("Very good. Mm hm") as one student continued to contribute. This short segment was precipitated by a teacher question and extended by another, but the teacher's TEvalL allowed for student contributions to be incorporated into retelling of a story.

Evaluations as affirmations did not tend to cluster together in most of the classrooms, but were consistently used across classes that used high-rates of TEvalLs. Teachers who used low-rates of TEvalLs show more inconsistent amounts of TEvalLs over their three observed lessons. Only three teachers showed strings of TEvalL over multiple turns of talk. These instructional episodes used a form-focused worksheet and the two other were conducting text talks (one in a small group setting, the other with the whole class) that had the students contributing their thoughts about poem and recounting events in an abridged text.

Low-quality evaluations as building blocks for extending the topic.

The third and final theme generated from multiple reads of this talk moves across lessons was their function as affirmations of a correct answer which was immediately built upon to further the topic of talk. Table 5.16 includes some typical examples of how TEvalLs were incorporated into later talk moves.

Table 5.16 *Typical Examples of Low-Quality Evaluations as Building Blocks for Talk*

Teacher	Speaker	**Discourse and Low-quality Evaluations as Building Blocks** (in bold)
22	Teacher:	…Who can give me another homophone without looking at the charts? Javier? O-N-E is the number. Can you name the other spelling?
	Student:	[*inaudible*]
	Teacher:	**W-O-N. Right. Very good. Won**. This is a number, and what does this ['won'] mean? Like the blank won the football game yesterday afternoon. Who knows who won the football game? Theo?
	Student:	I have absolutely no idea.
	Teacher:	**OK**, Adam? [*overlapping voices*] **Steelers, very good.** (P16:8, 1/24/11)
82	Teacher:	**Okay. Small animals can do big things.** Let's make sure we get the main part. "Sometimes stories about animals have a moral or lesson." Okay.
	Student:	Again you said what?
	Teacher:	Small animals can do big things.
	Student:	**Okay.** And you said?
	Teacher:	Don't tease someone.
	Student:	**Okay. Don't tease.** Would you say both of those are lessons?
	Teacher:	Yeah.
	Student:	**Okay. Good.** Don't judge a book by its cover because the turtle was small and slow, but the hare—and the mouse was small and weak, but he freed a lion. The tortoise was small like [*inaudible*], but he beat the hare. (P19:35, 3/3/11)
70	Teacher:	What might they do on their balconies?
	Student:	Sit outside.
	Teacher:	**OK. People sitting outside.**
	Student:	They could be cooking.
	Teacher:	**They could cook on their balconies.** They'd cook on their what?
	Student:	Grill.
	Teacher:	**Yeah.**
	Student:	My mother always [*unintelligible*].
	Student:	We don't even have a grill.
	Student:	I have a grill [*unintelligible*]. At least we go to the store to get it. We get more food.
	Teacher:	OK. [The setting is an] Apartment building in the summer. OK? Some things we may see, some things we may hear. This is the setting of? (P115:257, 2/9/11)

Note. Words in bold represent coded instances of low-quality teacher evaluations

An example of a common low-quality evaluation is captured by the pedagogical talk of Teacher 22 above ("Won. Very good. Won"). She accepted the student's correct answer and then

built on it with an uptake question. She then immediately scaffolded this question by suggesting a fill-in-the-blank statement, and then built on that to connect to students' presumed knowledge about the latest NFL fame. When a student answered this latter question, Teacher 22 used another TEvalL to praise his contribution, and then moved on to procedural issues in the lesson (not included in the excerpt). A similar pattern was clear in Teacher 82 focus on poetry and morals ("Okay. Don't tease. Would you say both of those are lessons?"). Both teachers used their evaluations to affirm their students' ideas and keep the lesson moving in a teacher-student-teacher pattern of talk, with different outcomes in the amount of student talk.

In a small reading group directed by Teacher 70, students were asked to think about a setting in relation to the story they were reading. The responses to this question generated affirmative evaluations from the teacher and more student talk where students responded to each other. However, this talk quickly became off-topic and the teacher stepped in to keep the talk focused on text comprehension.

In these examples, TEvalLs are brief affirmations of student contributions that teachers build upon to keep the pedagogical talk moving along. In some cases, teachers skipped TEvalLs and instead substituted follow-up questions in response to student contributions about texts to keep the lesson moving forward.

Teacher 23: …The Magic School Bus. Joanna Cole is the author of all the *Magic School Bus*. So she was born, someone look at the page and tell me when she was born?
Student: August 1944.
Teacher: **Where, where was she born not when, where was she born? Where was she born?**
Student: In New Jersey?
Teacher: Newark, New Jersey. And where does she live right now?
Student: Sandy Hook.
Teacher: **And what state is that?**
Student: Connecticut.

Teacher: Connecticut. So she was born in Newark, New Jersey and she lives now in Sandy Hook, Connecticut. (P77;79, 2/15/11)

This pattern, despite the lack of many low-quality evaluations, still demonstrated an acceptance of student responses and a similar stance toward the texts. Teacher questions and evaluations kept the discourse moving along with brief student answers accepted into the discourse through repetitions, explicit praise, or questions that built on the answer to form 'strings' of talk that moved through the texts with recall or report and questions requiring some inferencing. These strings functioned to have students attend to the information in the texts. In some cases, TEvalLs were observed to affirm extended talk that addressed a topic in more depth.

Low-quality evaluations as conditional affirmations.

Conditional affirmations were much more infrequent than affirmations across the lessons. Teachers rarely corrected students explicitly. Instead, they used affirmations to accept the answer but reformulated it toward a more acceptable answer. In one rare instance across these transcripts, an erroneous answer stops the pattern of talk and the teacher disagrees with the student.

Teacher 70: Why was [the character] kind?
Student: Because she gave Caroline and her family a feast.
Teacher: OK. She invited them and treated them to a feast.
Student: She's generous?
Teacher: **She's generous because of that.**
Student: Thoughtful?
Teacher: **She's thoughtful.**
Student: Considerate?
Teacher: **She's considerate.**
Student: Friendly?
Teacher: **She's friendly.**
Student: Royal?
Teacher: **Royal would describe her**.
Student: Trustworthy?
Teacher: **Mm. Not the best one to describe her.** But she's very generous. She's thoughtful, polite, kind, friendly. All those words would describe her based on what she did for the family. She invited them in for a feast and treated them like

royalty. So that's very kind and thoughtful of her to do that. Very good. (P115:315, 2/9/11)

While Teacher 70 disagrees with the student ("Mm, not the best one to describe her"), the teacher does not provide a reason why the character would not be trustworthy. Instead, she downplays the off-target answer and provides a high-quality evaluation by summarizing the previous (correct) contributions and cites evidence from the text. Teacher 63 also exhibited this strategy of providing an affirmation that became conditional, and then provided an explanation of the erroneous contribution.

Teacher 63: Now, it's not connected any more. There is no longer a land bridge, so they connected and they walked, and they traveled, and they ended up here, okay?
Student: And then they died?
Teacher: **Okay, so there was a land bridge. Well, then they had—they had their civilization. They didn't die.** We don't know what happened to them yet. We don't know. (P21:124, 11/30, 10)

These rare examples of TEvalLs as conditional affirmations point to the use of low-quality evaluations mainly as affirmations and not negations or corrections. Teacher seemed to avoid direct corrections in favor of a more indirect language use of conditionals and hedges, making low-quality evaluations function as affirmations in all cases.

Case summary: General patterns of low-quality evaluations.

Overall, classrooms with high rates of low-quality evaluations used TEvalLs to offer simple acceptances of student contributions, to offer conditional affirmations, or to affirm and move the talk pattern along. High-rate classrooms showed consistency of used across all of their observations. The most pervasive pattern after these was the focus on text comprehension in most instructional episodes, accomplished through a question-and-answer pattern of discourse. Simple affirmations with or without subsequent questions kept the pace of these instructional

episodes moving quickly through the content without the teachers or students dwelling on a single questions for more than a few turns in the vast majority of episodes.

These talk moves were the most widely applied code to these transcript data, and indicated the pervasiveness of a teacher-managed, question and answer talk genre that structured the majority of these language arts lessons. However, there was evidence for TEvalL acting in a dialogic capacity. There were about 100 more student explanations coded in classrooms with high-rates of TEvalLs than classrooms with low-rates, and many of these explanations co-occurred with low-quality evaluations, among other talk moves.

Case 4: Student Explanations.

Student explanations were hypothesized to be indicators of dialogic instruction since reciprocal dialogue in a classroom necessarily means student participation. Reading interventions that emphasize discussions around texts are undergirded by the idea of literacy as language development. Language development not only needs rich input, but learners need to engage with others to develop their language and conceptual understandings (Swain, 2005; Vygotsky, 1986; Wells, 1999). Language development can be supported through adult support, such as adjusting speech to accommodate their estimation of a child's understand of a word (Weizman & Snow, 2001), or encourage topic elaboration for students to produce "more complex structures and create opportunities for adults to [tailor support]" (Dickinson & Porche, 2011, p. 871). These understandings implicate the need for instructional talk to shift from teacher-directed to shared discourse between students and teachers, necessarily increasing student talk during instruction. In the present study, however, extended turns of student talk had a negative relationship with student reading comprehension, indicating that the higher frequency of student explanations a student experienced during instruction, the lower their reading score was predicted to be.

Student explanations were utterances that contained two or more ideas and spanned at least two lines of transcript. When students shared their writing, this was also coded as student explanations, just in a written mode of communication. These codes were not as frequently applied to the transcript data as question and evaluation codes. Student explanations accounted for 4% of the total codes applied to these data. The frequency of student explanations varied widely across classrooms, and averaged 0.16 explanations per minute, or about 9 an hour. Student explanations ranged from one sentence with two ideas ("So if it said she found it at the bottom, you would know that she would be looking through the whole box") to 20 sentences of uninterrupted speech. The lengthiest explanations were all student written texts that were shared aloud.

Across both sites and across all grade levels, the content and function of student explanations fell into two general themes: explanations 1) addressed language arts content, usually in response to teacher questions, and 2) shared personal experiences, background knowledge, and opinions.

Student explanations as instructional content.

The majority of instances of extended student talk were in response to teacher questions about the content of texts. A smaller number were explanations about vocabulary, again in response to teacher queries. Student explanations address the content by providing speculations and inferences about narrative texts, detailed students' reasoning and opinions about texts, recounted or summarized plot events from a story, or defined meanings of events in texts or vocabulary words. The follow Table 5.17 presents typical examples of student explanations providing these sorts of content knowledge as part of the instructional discourse.

Table 5.17 *Typical Examples of Student Explanations as Content*

Teacher	**Speaker**	**Discourse and Student Explanations** (in bold)

27	Teacher:	That was his nickname. Gordie was his nickname. Jason, what do you think?
	Student:	**Well, I think it's because at first he didn't realize the severity of the situation, and so he just kept calling him Gordie. Now he realizes how serious it is. When my mom gets really serious or gets really, really mad at my little brother, she calls him [student name], his whole name**.
	Teacher:	Oh, when you get all three names. I like how you said the severity of the situation. So he was still kinda Gordie, Gordie, but then he realized he's getting close to his fear place, and he realizes how severe the situation is. Good. Yes?
	Student:	**I have an add-on to what Jason said. Sometimes when you're in a situation, it's all fun and games, but then you see what this really—then you finally come to, oh, this is serious. You have a total mind change.** (P3:85, 3/3/11)
10	Student:	**That girl Guapa, she went to a house and the other ones told the queen to talk to her. Then the queen said don't to her don't go, because then they will know their secrets where they're living. Doctors will put them in cages and study them, and put them in land.**
	Teacher:	Uh-huh, wow, good. You really know what's going on. So they're trying to keep this a secret, so that they don't get questioned or studied about this, right? (P6:150, 1/31/11)
30	Student:	**Miss [Teacher 30], he didn't really persuade Matt to let him in because Matt was just thinking of what his Dad said, of being polite to strangers. And not lured him in. He was like—**
	Teacher:	Yeah, I think, I think Ben tried to be persuasive, but then of course Matt does, Matt does think about what his father would do and how you're supposed to be polite and hospitable.
	Student:	And Matt told him to come in, and he, and he got his food—
	Teacher:	Okay, wait, wait, wait. Shh. Don't go ahead. What, what does he, what actually happens here, Javier? (P62:57, 2/9/11)
63	Teacher:	What could have happened?
	Student:	**Other people came, took other people, took other people and they stored everything away.**
	Teacher:	Oh, okay, so it's another set of people come in and all of a sudden no one else is there. So what did they—what did those new people come in and do to the Maya people? (P21:163, 11/30/10)
73	Teacher;	He's not definitely the most beautiful and you're telling me he shouldn't be saying that? Why?
	Student:	**Because sometimes you can have—the tree he got—his leaves died. In the spring he got them back.**
	Teacher:	Why did the owl do this to him? What did the owl want him to do? (P19:41, 2/9/11)

Note. Words in bold represent coded instances of student explanations

In a lesson characterized by a lot of student explanations, Teacher 27 guided a whole-class talk about a character's dilemma in an unabridged narrative text. In much of the instructional episode, Teacher 27 solicited student thoughts about the problem and the character's thinking. In the excerpt above, the student explained his opinion of the fictional situation, and then supported his view with an example from his own life. After the teacher's affirmative evaluation, another student contributed along the same lines, suggesting that his peer's contribution spurred him to generalize about the situation ("I have an add-on to what Jason just said," and "Sometimes when you're in a situation…"). Throughout this lesson, Teacher 27 tolerated many personal examples related to the events in the story.

The excerpt from Teacher 30's class was a typical example of a student addressing the content of her question about the events in the chapter book, *Sign of the Beaver* (Speare, 1983). This explanation contributed to the pedagogical talk by its accuracy and importance to the teacher's focus on the character's motives, as indicated by her positive evaluation and restatement of the student's explanation ("Yeah, I think, I think Ben tried to be persuasive, but then… Matt does think about what his father would do…"). However, when the student attempted to continue explaining the plot line, the teacher asserted her control over the talk to manage the chronology of the retelling. Student explanations that addressed content accurately worked in similar ways in lessons, becoming part of the content of the lesson while still being framed by teacher-managed episodes. The one example of a student explanation in the only SEI class in the sample worked in a similar way. In the excerpt on Table 5.17, Teacher 10 posed a question about the events in a story, to which a student provided an accurate and detailed summary of the key events. The teacher provided a positive evaluation, and paraphrased the

student's response for the class, similar to the tendency of other teachers across the sample and grade levels.

All student explanations that addressed content in response to teacher questions were a few lines of transcript at most. Teachers whose instruction included frequent student explanations tended to consistently use them in all three observations throughout the school year. The student explanations with Teachers 63 and 73, while frequent, were characterized by short responses to questions about text. However, many of these explanations lacked specificity ("other people came, took other people…"), to which Teacher 63 followed up by providing more clarity for the class ("so it's another set of people…"). Teacher 73 followed an unclear explanation with two uptake questions soliciting more information, to which the student again responded with a lack of specificity, and Teacher 73 moved on to another student (not included in excerpt above). This happened frequently across classrooms where student explanations were not clearly articulated or specific enough for the teacher's purposes of text comprehension. In general, student explanations were followed by teacher agreement, rephrasing, or uptake questions.

When students' contributed off-topic utterances, the resulting teacher talk varied. During a small group text talk on the meaning of a poem, a student provided an explanation that she noted was off-topic but continued to share. Teacher 73 followed this with a simple evaluation and moved on, seeking an answer to her question.

Teacher 73: Have you ever seen somebody's eyes who were grayish-blueish-green?
Student: I have blue.
Teacher: No, not one of them, all three.
Student: **Well, I know something. Not so much like this, but usually when my dad takes a picture of us, some of us—on his phone and some light hits them and then the camera flashes, the light sort of catches our eyes and then it just gets caught in the picture.**
Teacher: Good. Ashanti? (P20:58, 3/3/11)

In contrast, when a student provided an ambiguous explanation of the difference between 'sympathy' and 'empathy,' Teacher 16 pushes the student to explain her thinking. In this excerpt, the student is held accountable for her explanation of the two words, which presumably worked to benefit the class' understanding of them as well.

Teacher 16: So how is sympathy different from empathy? Are they connected or are they not connected?
Student: **They're connected in one way but in others--they're like cousins--they're like brothers in one way and long-lost cousins in another way.**
Teacher: Okay, in what way are they like cousins?
Student: One of them is meaning, like, you could imagine yourself being this, doing this.
Teacher: Okay, so which one would be that--that you can actually imagine yourself doing that?
Student: Well, not doing that but you feel like that.
Teacher: Or feeling that, okay. Which word would that be, empathy or sympathy? Empathy. Do you agree?
Student: **Yes, and when they're close, you can imagine you were really in their shoes. You were really wearing their shoes.**
Teacher: Well, that's empathy. (P82:18, 1/31/11)

The majority of student explanations stayed on-task to address content by defining words, providing details of their reasoning or speculations, or summarized texts, and were followed up by teachers' evaluations, restatements or uptake questions. In this way, the content of students' extended responses provided the class with content in line with teachers' focus of the lesson. These explanations were incorporated like a teacher explanations, but managed by the teacher through her restatements or uptake moves. However, none of these student explanations were longer than three sentences or three lines of transcript. The "extended" nature of these explanations was much briefer than teacher explanations.

Student explanations as shared experiences.

Some teachers were observed to structure time in their lessons for students to share their experiences or their work with their peers. By having students explain their own experiences

publically, these personal experiences became shared experiences of sorts because they were part of the instructional discourse. These student explanations tended to cluster together in certain transcripts because teachers organized certain times to allow as many students to share as was possible during the episode. For this reason, there were few follow up moves after these kinds of extended student talk moves. Some common examples are presented on Table 5.18.

Table 5.18 *Typical Examples of Student Explanations as Shared Experiences*

Teacher	Speaker	Discourse and Student Explanations as Shared Personal Experiences
83	Teacher:	Yeah I don't know. I don't know any part of the boat. I don't really like going on boats. It makes me seasick. I don't really like being on the water.
	Student:	**When we were riding we even got to put our feet in the water and our hands in the water. One time when we were riding I was hanging off the boat and the boat got tipped like this and my hand went all the way under the boat, for like 30 seconds it was fun.**
	Teacher:	Sounds traumatizing. Yeah, Keyon, last one. (P33:80, 2/4/11)
18	Student:	Snore, snore, snore. That's all that my pet does. Throw a stick, it just lies there. Call his name, it just picks up its head and moves it side to side. He just sleeps most of the day and can never get excited. (P79:133, 2/26/11)
81	Teacher:	Teacher 81: Before you're done, Tierra, can you just play one or two songs and explain why you picked those titles?
	Student:	**I picked "Rocking From Mother to Father," because she lived with her mother for most of her life in Milwaukee, and then she lived with her father in Nashville, Tennessee. I picked "Oprah's Gold Medal for Reading," because she won a gold medal for the Reading Foundation.** (P108:186, 3/3/11)
81	Teacher:	[*at end of learning activity*] Okay. Celebrations and reflections on that rotation. Brianna?
	Student:	**Me and Jackie, we was reading-- we read this story, and we did great on it and** [*cross talk*] **was trying to interrupt us, we kept on reading.** (P106:98, 2/3/11)

Note. Words in bold represent coded instances of student explanations

Teacher 83 had an instructional episode characterized by her elicitation of personal examples loosely tied to the theme of the text. Student explanations received some brief comments from the teacher, but it is clear Teacher 83 intended to move through students to share

without much taking up of their experiences ("Sounds traumatizing. Yeah, Keyon, last one"). This pattern of emphasizing the sharing personal examples equally by students was corroborated in other lessons, where teachers would note the appropriate time for sharing. For example, after a student fit in a brief experience during a Q & A about a reading, Teacher 25 responded by saying "Yeah, that's scary. He should have been wearing a helmet. Okay, we're not going to share stories right now. I'll let you share stories later" (P42:136, 2/15/11). Unlike student explanations as content, in many cases these types of explanations were not seen as content of the lesson.

In contrast, Teacher 18 had students write short texts that each shared to the class in order for their peers to infer the main idea of the writing (which in this case is 'my pet is lazy'). There was more teacher and peer feedback on students' writing, but the activity was structured to move quickly through student work so that all had the opportunity to share. In similar fashion, Teacher 81 used student presentations to provide opportunities for extended student talk. In the excerpts on Table 5.18, students were held accountable for their explanations through follow up moves that required the student to provide more required information, or by the implementation of regular sharing experiences at the end of learning episodes, a routine that had students reflecting on the work they accomplished. However, there was no following extended talk as presentations were followed by another.

Student explanations that shared personal experiences or writing assignments tended to cluster together in certain episodes dedicated to having students share, and while these episodes drew from a text theme, many were not explicitly connected back to the content of the class. These episodes seemed more to be with providing time for students to talk, and moving the instructional discourse quickly so most students could share. Many instances of this type of

student explanation were longer than the average three lines of transcript used by students providing instructional content explanations, discussed in the previous section.

Case summary: General patterns of student explanations.

Both themes describing student explanations reveal a lack of depth of both talk and content knowledge when these explanations were incorporated into the lesson discourse. Student explanations as content did demonstrate student attention to the lesson as they were routinely responses to teacher questions about texts, but for 'extended' talk, these utterances were consistently brief, no more than a few sentences or lines of transcript which barely met the minimum of the code definition. Many were also unclear or somewhat off-topic, which was met with varying degrees of correction in follow up moves. In these utterances, it is a possibility that this student-derived content may have acted to confuse or mislead peers from the teacher's focus. Student explanations as shared experiences were lengthier but less frequent and almost always contained within instructional episodes devoted to allowing students to share their experiences or writing. Since the instructional emphasis during these episodes was on providing all students with the opportunity to share, there were few instances of the content of these explanations being taken up for in-depth exploration or explicitly connected to texts of reading instruction. This may have attenuated the efficacy of exposure to peer language across these lessons.

Case 5: High-quality Teacher Questions.

High-quality questions were any type of question that elicited the cognitive skills of inferencing, predicting, evaluating, generalizing, or analyzing information (usually from texts). Analyzing the cognitive quality of pedagogical questions can be imprecise in research (Cazden, 1988), and the dichotomous coding of questions in these data (high- vs. low-quality questions) is a simplistic measure to assess the thinking skills the questions elicit. However, the coding

scheme in the present study was created to look at the correlations between discourse moves and reading outcomes, and not at discrete cognitive skills and their relationship to reading. The cognitive skills implied by teacher questions were only coded to provide a sense of how teachers may be eliciting student talk and reciprocal talk between participants. Identifying the cognitive levels of questions can give us a sense of the teacher's stance toward texts and reading: whether s/he treats reading as a reporting or recall activity, or as an activity that incorporates the "possibility of the speaker's … perspective and particularity" (Nystrand et al., 2003, p. 147) through the stimulation of discussion. Following the work of Nystrand and colleagues, high-quality questions are similarly defined in the present study.

In the final HLM model in Chapter Four, high-quality questions were shown to be predictive of lower reading scores on the WMLS Passage Comprehension reading assessment. This finding contradicts the hypothesis that these questions can elicit dialogic instruction, and contradicts empirical literature suggesting that these higher level thinking skills have been related to better reading outcomes (Nystrand et al., 2003), resting on the argument that opportunities to engage in these sorts of thinking skills, skills used by good readers, help support students' reading. However, the discussion in Chapter 6 problematizes the quantitative finding.

Across all lessons, there were 1574 instances of high-quality questions, which accounted for 7% of the total codes applied to the data. High-quality questions averaged almost 20 per hour of instruction. (In contrast, there was an average of 44 low-quality questions posed per hour). There were no significant differences in the rates of high-quality questions used between the three grades or by site.

In refining the inductive coding of high-quality questions across the 31 classrooms, two themes emerged. The first theme addresses the content of these questions: an overwhelming

majority of high-quality questions required students to infer and speculate about text, and very rarely to analyze, evaluate, or generalize about text. In conjunction with this feature of high-quality questions was the pervasive efferent stance toward texts taken by all teachers in this study. A second theme notes the pattern of talk around these questions. High-quality questions that occurred in classrooms with high rates and low rates of use tended to be scaffolded. These speculative and inference questions were scaffolded to lower quality questions that required students to recall known information or identify information explicitly stated in the text.

High-quality requests for inferences.

The most salient pattern to the content of high-quality questions across the 31 classrooms was the efferent stance toward text that they reflected. This was made clear by the prevalent use of these questions to elicit inferences and speculations about texts. Though the high-quality question code was applied to all questions that asked students to use skills of inference, evaluation, analysis, or generalization, the overwhelming majority of these questions gave students opportunities in the one skill of inferencing (predicting a character's next actions or speculating on a theme of an unread story draws on the ability to use text clues to infer meaning, so I consider these foci as one skill here). Table 5.19 presents some typical examples of high-quality questions that had students engage in these cognitive skills.

Table 5.19 *Typical Examples of High-quality Questions*

Teacher	Speaker	Discourse and High-quality Questions

27	Teacher:	And he was dismayed that he was thirsty again. What is dismayed? **If you're not sure, can you infer or make an educated guess about dismayed?** Jason?
	Student:	It's sort of like he was sad about it, like [*unintelligible*] ooh.
	Teacher:	Yeah, so you were in distress. You're troubled about it. **Can we infer one last inference about Doug and his feelings for his brother?** Just based on that last paragraph. At this rate, there will be little left for Gordy, so he took the cap off his canteen and drank only a swallow. What can you infer about his feelings for his brother?
	Student:	That he cares about him because he drank most of it, half of it?
	Teacher:	He only had a swallow. Just a little [*unintelligible*]. So he did care for him. Instead of saying 'I'm just going to chug this whole thing, I don't care about' [noise obscures]. He was thinking ahead, and he was worried about him. **Would you do the same thing for one of your brothers?**
	Student:	
	Teacher:	Depends.
	Student:	[*laughter*] It depends on which brother or the situation. OK, yes.
	Teacher:	Is Gordy older or younger?
	Students:	**Do you think Gordy is older or younger?** I think he's younger…Younger…. Older [*overlapping voices*] [P2:213, 2/16/11]
62	Teacher:	(*stops reading aloud*) **What do you think Brer Rabbit is going to do right now**? Cindy, what do you think?
	Student:	Try to get Hippo to play tug of war with him.
	Teacher:	With him, you think? **Do you think Brer Rabbit would win tug of war against Hippo?**
	Student:	No.
	Teacher:	**So what do you think he's trying to do, Keon?**
	Student:	[*unintelligible*]
	Teacher:	He's not moving slow. He's excited about this. He wants to get this competition on. …
	Teacher:	OK, that's a good idea. And we'll take one more. **Is there anyone else who I haven't heard from this story who would like to tell us what they think?** Michael? [10:43, 12/2/10]
16	Student:	[*Response inaudible*]
	Teacher:	That's great! Can you give me an example? To analyze something is to explain something. What do you mean by that?
	Student:	[*Response inaudible*]
	Teacher:	Ok, in a math problem you explain your answer. **How do you explain something in a math problem?** What are you doing? What's the process?
	Student:	*[Response inaudible]*
	Teacher:	Ok, so in a math problem when you are asked to analyze your answer, you're going to explain it by drawing a picture is one example. [*another student name*]? [P81:55, 12/15/10]

79	Teacher:	What do you think without looking at the answers? Let's just think on our own. **Why do you think the duckling would want to run away?**
	Student:	Because the others didn't want him around
	Teacher:	Because the others didn't want him around? Okay. Gina?
	Student:	Because the other were chasing him and pecking him and they were saying bad things about him.
	Teacher:	They were. They were chasing and bothering him, being mean to him. Yeah, they didn't want him around, right? They were being very mean. What's the first answer there for number four? Was the duck upset with his mother? No. Do you think he wanted to live away from all the other ducks?
	Student:	No. (P104:44, 2/25/11)
70	Teacher:	So how do you think the kids feel about their mom? **How do you think**
	Student:	**the kids feel about their mom?**
	Teacher:	That she like, is sort of mean?
	Student:	**Why?**
	Teacher:	Because she won't let them get a pet, but she wants to go other places. Okay. So, she'll – **you think she wants to do things for herself but not for her family? Are you thinking that? Or do you think she's**
	Student:	**looking out for her family?** I'm not saying that, but she's trying to clean up around the house, but she wants the money to go out somewhere and get a break, but she doesn't want their kids to have a pet. [P118:218, 2/14/11]

Note. Words in bold represent coded instances of high-quality questions

Teacher 27 consistently used high numbers of high-quality question across her three observations, spending a lot of time asking students for inferences based on three different narrative texts. She used these questions in the contexts of text comprehension and reading response writing similar to the writing on the state's standardized tests in ELA. Near the end of this excerpt of a read aloud, the teacher posed a more evaluative question, asking students to evaluate their own experience of worrying for a family member for its applicability to the current learning task. She then took up a spontaneous student question to engage students in more inferences of the characters in the text. Similar to the inference questions posed by Teacher 27, in the excerpt from Teacher 62, the students were asked a series of 'what do you think?' speculative questions on the plot line or characters' behaviors, and in both classes, student answers were very brief.

Teacher 79 worked with a multiple choice text format to elicit inferences to correctly answer the test question. This example clearly indicates the typical cognitive skills elicited by the typical testing materials used in these classrooms. Mirroring this test emphasis, the majority of the high-level questions asks students for inferences and predictions about texts, of which the majority were narrative.

One of the few high-quality questions that had students analyze, Teacher 16 first used two uptake questions that asked the student to draw on her background knowledge ("Can you give me an example? To analyze something is to explain something. What do you mean by that?"). The teacher then pressed the student to explain how one would explain something in math. This seemed to be asking the student to analyze how the genre of a math explanation gets accomplished, which was asking the student to think more metalinguistically about what the genre of explanation entails.

Finally, Teacher 70 provided multiple examples of questions eliciting more inferences of text from her students. Her "why?" question provided one student the opportunity for an extended explanation, that she then clarifies after Teacher 70 attempts to verify her response ("Are you thinking that?"). While still in an IRE pattern of discourse, the student is able to respond and clarify her thinking to her small reading group.

This IRE pattern of discourse was reflected in the high frequency of use of the Q & A talk genre (followed by read alouds and independent work) that characterized the instructional episodes where most high-quality questions occurred. Another pattern to these questions was that they were most frequently coded within instructional episodes that used unabridged and testing materials, and within episodes that were focused on text comprehension, which is no surprise considering the vast majority of these questions required students to speculate on and infer from

texts. And like uptake questions, high-quality questions did not cluster together. There were no identifiable patterns where high-quality questions or other talk moves begot high-quality questions.

While the majority of these questions practiced students' skills of inference, there was evidence of other cognitive skills elicited through these questions. A somewhat common questioning pattern was to ask students for examples on a theme connected to a text. These questions seemed to be mainly asking students to generalize information. For instance, Teacher 27 began a lesson by asking "What are some ways that people might deal with being afraid?" to which students generated examples of fears and how people tend to react to them for almost 18 minutes of instruction. Another asked if students could "think of a time when [one] might scowl at someone" during a vocabulary lesson. Both questions required students to think about the generalities of these situations, drawing on their background knowledge. This was later used to apply to their understanding of the text. There was a moderate trend for classrooms characterized by low rates of high-quality question use that showed more variety of cognitive skills than the high-rate classrooms, including the only questions that elicited synthesizing information in the entire sample ("What do all these words that we just acted out, what do they have in common?" [P111:33, 1/28/11] and "What was the connection between that?" [P 87:33, 2/15/11]).

The preponderance of high-quality questions eliciting inferences indicate an efferent stance toward the readings, constructing the activity of reading as a skill to glean information from text (Rosenblatt, 1994). There were only two questions in the case study of high-quality questions that could be construed as having students respond to a text from a more aesthetic perspective of reading: Teacher 73 paused during her many questions requiring students to infer the meaning of poem to say, "Oh, that's interesting. Wait, did this author do it again? What do

you notice about the colors?" (P20:211, 3/3/11), temporarily focusing student attention on the author's choice of words. This pervasive efferent stance toward reading was accompanied by the lack of diversity in cognitive skills required of students when engaging with the meaning of texts, as well as a lack of extended student talk in response to these questions. Using Atlas ti qualitative software, co-occurrence searches resulted in only 17 times across all 88 transcripts where student explanations followed high-quality questions. This pattern is discussed further in the second thematic feature of this case.

Scaffolding from high- to low-quality questions.

A second theme generated from multiple reads of transcripts was the frequent way high-quality questions were scaffolded by teachers toward more recall or recitation questions. Many of these high-quality questions were posed and then scaffolded to easier questions like either/or, recall or known-answer questions, or in some cases, prompting students with hints at an answer the teacher deemed acceptable. Typical examples of high-quality question scaffolds are included on Table 5.20.

Table 5.20 *High-quality Questions and their Scaffolds*

Teacher	**Speaker**	**Discourse and High-quality Questions** (high-quality questions in bold, scaffolds in italics)
70	Teacher:	**So if the heat would be on, what would be happening?** *Would the windows be open? The windows would be ---?*
	Student:	Closed.
	Teacher:	Right. [P15:28, 2/9/11]
10	Student:	I think [the text is] going to be about a girl who has a secret.
	Teacher:	Maybe a girl who has a secret. **What tells you that?**
	Student:	Because there is a girl and this is the secret.
	Teacher:	You see a picture of a girl, and *what's the title?*
	Student:	The Secret Footprints. [P6:44, 1/31/11]
16	Teacher	Maggie does not want to write cursive. **Why do you think she does not want to write cursive?** *Is it hard? Maybe it's hard for her. That's one reason.* What's another reason? [P81:149, 12/15/10]

83	Student:	Because they haven't talked to each other.
	Teacher:	They haven't talked to each other for the whole day, so they have a whole lot to say. So how is that different from if they were allowed to talk to each other all day, and then they get to this last period class, and the teacher says, OK, you can only write to each other? **What might, how might their writing look different?**
	Student	[*unintelligible*]
	Student:	Because when you talk, people have something to say, and they run out of words. So when you're writing, it might look like a lot of words that you're writing.
	Teacher:	So it might be a lot? **Do you think they're going to have a lot to say to each other?** *They had been able to talk all day long. Are they going to have a lot to say in writing?* (P29:130, 12/10/10)
25	Teacher:	Okay, just one more thing I want us to practice before we stop. What could we write for a topic sentence? Because that's one of the hardest things for this classroom to do, for this class to do. **What could we write for a topic sentence?**
	Student:	A topic sentence.
	Teacher:	**If we were to start writing this open response what could we write for our topic sentence?** *What is the strategy I've taught you over and over again, what's it called? Like baseball*.
	Student:	That's closing sentence.
	Student:	Steal and slide.
	Teacher:	Steal and slide, steal part of the question and slide it into your answer. **What could we write, Madison, what part of the question do you think we should steal?**
	Student:	Why Michael's bike would be unsuitable for speed. (P42:301, 2/15/11)

Note. Words in bold represent coded instances of high-quality questions; Words in italics highlight scaffolds for these questions.

In the first three excerpts from Teachers 70, 10, and 16, the original high-quality questions were scaffolded by a cloze question or a test question ("What's the title?"), and by Teacher 16 answering her own question. These scaffolds give students very little to do in these talk segments. Each student answer after these scaffolds did not qualify for a student explanation code.

Teacher 83 picked up on a student contribution and after some description of the text plot, she posed a high-quality question that asked students to analyze the possibility of an opposite scenario from the text. This question elicited an extended response from one student, a rarity, which was then picked up by the teacher to pose another high-quality question eliciting a speculation from

her students. This second question was heavily scaffolded, at first by essentially giving an answer in her earlier utterance ("They haven't talked to each other for the whole day, so they have a whole lot to say"), and again by recounting the plot briefly ("They had been able to talk all day long. Are they going to have a lot to say in writing?"). The second high-quality question ("Do you think they're going to have a lot to say to each other?") was scaffolded down to a question that elicited more of a yes/no answer from the student, but keeping the students on track with the teacher's attempt make sure they comprehend of the significance of the plot event.

Teacher 25 also showed the pervasive scaffolding of high-quality questions that most teachers used. In her first question, she asked the students to write a topic sentence for a reading response using test-like materials. This required students to generalize from their understanding of the story in order to construct a topic sentence that captured the prompt and the main idea of the student's response. After an off-target answer, Teacher 25 scaffolded this question by posing a recall question, reminding students of a previous lesson. This elicited the answer she was looking for, and provided a platform for her to repeat the high-quality question with more focus ("What part of the question do you think we should steal?"), to which a student responded with an acceptable suggestion.

A counterintuitive finding of the function of these high-quality questions was that very few were followed by extended student talk. Less than 1% of high-quality questions were followed by the student explanation code. Like uptake questions in the previous case study, teachers across the classrooms and across the grade levels seemed satisfied that students provide short phrases as answers to their questions. The IRE pattern of instruction, reflected in the frequency of use of the Q & A talk genre, was so pervasive in these classrooms that it would seem that no matter the question, teachers were willing to accept short answers, and students did not offer extended responses, thus continuing this pattern.

The frequent use of scaffolding high-quality questions may have mitigated student talk. There was a consistent pattern of scaffolding high-quality questions to more recitative questions in high-rate and low-rate classrooms. The following two segments of talk from Teacher 62 are representative of the kinds of scaffolding in the low-rate group of classrooms, and the smaller amount of student talk that can be generated in this pattern of questioning.

Teacher 62: Alright. Celine, what can you tell me? **What inference did you make?**
Student: That the snow is cold outside. The snow, it melted when the sun came out.
Teacher: Okay, but that's something that we talked about from information in our heads. Okay? This little boy he says that he made a little pillow for his snowball. It was gonna sleep with him and then it ran away, but first it wet the bed. What I want to know, **what really happened?** *What do we know about snow that maybe this little boy didn't know about?* Lydia?
Student: That it melts.
Teacher: Okay, so what happened to his snowball?
Student: It melted.
Teacher: His snowball didn't run away, and it didn't wet—
Student: The bed.
Teacher: —the bed. Kind of like what wet the bed means. It did what?
Student: Melt.
Student: It melted.
Student: It melted.
Teacher: It melted (P15:33, 2/10/11)

Teacher 62: [*repeating student response*] Biting and buzzing back. **Yeah, but what did [the character, a fly] do to get his—to achieve what he wanted to?** *What did he keep doing?* You need to give examples from the story. *Who was he buzzing and biting?* I don't know—I haven't read the story before. Who was he doing that to? You need to give every detail. Yeah, but honey he wasn't telling him to go, … *what did he keep doing to moose?*
Student: Biting and biting the moose.
Teacher: Yeah, you need to add that now. Keep going. He kept biting, no, no leave that period. You need to start a new sentence with a capital letter. There you go. (P12:115, 1/13/11)

In both examples, the teacher posed high-quality questions that asked students to infer meaning from texts, and then posed more restricted questions that solicited short phrases. In the first example, the first high-quality question without scaffolds received an extended response, which was deemed off-target. Teacher 62 asked her students again, but this time posed a more

specific and constrained question to get at the answer she was seeking ("What do we know about snow that maybe this little boy didn't know about?"). Responses to this were brief and accepted by the teacher. In the second excerpt, Teacher 62 counseled a student writing a reading response by first asking the student to speculate on how the character managed to achieve his goal. This was immediately scaffolded to two low-quality questions that asked the student to recount the actions of the character ("What does he keep doing?" and "Who was he buzzing and biting?"). This scaffold was in service of having the student better understand audience expectations for this genre of writing. As academic writing uses a distant and authoritative voice (Schleppegrell, 2006), the teacher was explicit here about the need to be specific for a reader without shared background knowledge.

One difference in high-quality questions between high-rate and low-rate classrooms was that many of the high-quality questions posed in low-rate classrooms were surrounded by a lot of teacher talk. For instance, Teacher 83 provided quite a bit of explanation after she received a short answer responding to her 'why' question:

> Student: She doesn't [*background noise*].
> Teacher 83: **Why?** You're right.
> Student: [*background noise*]
> Teacher: Okay, if she is a TT therapist and then suddenly a report comes out saying that it's not real, what's going to happen to her?
> Student: She'll get—
> Teacher: She's going to lose her job, so obviously those people who perform TT are obviously going to not be happy with these results. They're going to be against it because it could mean that they could lose their job or their—what vocabulary could we say, that they're going to lose their good what? I see a couple people now—I see more people though—a perfect example of how we can use our vocabulary. (P47:168, 2/1/11)

The teacher's first asked the student to explain her answer, and then scaffolded her next question by setting it up so that this uptake was a rather simple prediction ("…then suddenly a

report comes out saying that it's not real, what's going to happen to her?). In her third turn, the teacher interrupted the student to answer herself within an extended explanation that links back to the focus of the lesson, vocabulary. Similar in function to teacher explanations, this sort of pedagogical talk would seem to limit student opportunities to practice these skills while at the same time providing accurate content.

Case summary: General patterns of high-quality questions.

In this small case study of high-quality questions, two themes were identified in characterizing the patterns of use of this talk move. High-quality questions were hypothesized to support increases in student talk through increases in dialogic instruction, in turn leading to better reading outcomes. However, statistical findings indicated that this type of question in fact predicted lower reading outcomes. Both themes suggest possible reasons for a diminished impact of the dialogic potential of these questions. Qualitative explorations of these questions suggest that the vast majority of these high-quality questions functioned to provide students with opportunities to speculate and infer information from narrative texts within a question and answer talk pattern where teachers focused on student reading comprehension, which did not provide a variety of opportunities for students to generalize, analyze, or evaluate information. Secondly, in many instances across the sample, these speculative questions were scaffolded to lower quality questions that narrowed possible answers and generated little extended student talk. However, it is important to contextualize these themes. First, high-quality questions were highly correlated with uptake questions which predicted improved reading scores. Further, in these classrooms, it would seem that teachers providing frequent high-quality questions and their concurrent scaffolding were providing opportunities for students to engage in inferential

thinking, a skill show to explain variance in reading comprehension outcomes at these grade levels (Cain, Oakhill, & Bryant, 2004).

Conclusions

The five significant findings from the multilevel models suggest the importance of instructional talk for students' achievement on the reading comprehension measure, but how this might be so is complex. In this chapter, the five talk moves predicting student's scores were explored as separate cases to focus on each in some depth within the context of the lessons in the sample. Each case study analyzed the content and function of each talk move across all the classrooms. The intent of these analyses was to move beyond the quantitative findings that assumed the frequency of dialogic talk moves predicted reading achievement, to detailing the quality of features of dialogic instruction, and how they functioned to "get things done" (Stake, 2006, p. 1-2). Themes from these inductive analyses were derived through coding and comparative memoing. Patterns and variations within and across classrooms were identified using Atlas ti 6.2 (Scientific Software Development, 2011) qualitative software tools. An additional analysis looked at each talk move within classrooms grouped by the frequency of observed use, and consistency of use of the talk move by individual teachers.

In general, the five talk moves were found to function in similar ways to the definitions in the empirical literature, but in more attenuated ways. For instance, uptake questions were clearly built on a reciprocity between student ideas and teacher intent to assess student comprehension, but these questions mainly asked students to recall or infer information from texts; little analytic or evaluative thinking was elicited by the questions. High-quality questions were also mainly opportunities for students to infer or predict information from texts. However, in this case, these questions tended to be scaffolded to more attenuated questions, to which brief answers were

satisfactory. Similarly, teacher explanations frequently provided vocabulary definitions and description of concepts encountered in texts. Summarizing content, sharing personal experiences, explaining reading strategies, or conducting think alouds were not regularly observed. Both the scaffolding of high-quality questions and the frequent use of teacher explanations seemed to produce limited opportunities for extended student talk.

The preponderance of teacher-managed instruction and teacher talk, including questioning, indicated low levels of student language output, and a lack of depth to the output. There were few genuine text-based discussions, either planned or that spontaneously happened, to encourage more dialogue between students and teachers. One facet of instructional talk supporting this conclusion was the lack of a consistent pattern of student explanations following high-quality questions, which runs counter to the hypothesis that high-quality questions can encourage student talk through discussions (Nystrand et al., 2003). There were few instances of exploratory talk, conceived by Elizabeth and colleagues (2012) as the highest quality of discussion talk. Overall, dialogic instruction in these classes was circumscribed, mitigated by instructional talk genres, some pedagogical materials, and a pervasive stance toward texts as tools to extract information to be applied in narrow ways, such as answering textbook or practice test questions, or writing state practice test prompts. To be sure, there were some classes that consistently did not reflect this general pattern of dialogic instruction, using small group text talks, and exhibiting some instances of unstructured and somewhat exploratory discussions between students and teachers. In general, all classes showed some dialogic practices during instruction, which suggests the importance of identifying and testing those promising features that incorporate more dialogic practices into language arts instruction.

Chapter 6 – Integration of Analyses: Resulting Themes of Instruction

Language skill has consistently been related to reading comprehension (Cutting & Scarborough, 2006; Dickinson et al., 2003; Lawrence & Snow, 2011; Ouellette, 2006; RAND, 2002; Snow, 1991). The simple view of reading (Hoover & Gough, 1990) draws on the empirical evidence for this relationship to posit that linguistic comprehension skill is one of two key components of reading comprehension (the other being decoding skill). One crucial implication of this model of reading is that "instruction facilitating linguistic comprehension should … facilitate reading comprehension" (Hoover & Gough, 1990, p. 153). How teachers guide and support students in developing the language abilities to support their reading comprehension is still not fully understood, and the express intent of the current study is to identify mechanisms of productive discourse by quantitatively exploring their relationships to reading comprehension, and qualitatively to describe their content and function within naturalistic classroom discourse. This study hypothesizes that productive discourse is dialogic and reciprocal between teachers and students, in part to facilitate this linguistic comprehension skill. An additional impetus for this research is the current interest in instructional quality in educational policy in the US. The impact of pedagogical moves in literacy instruction has become increasingly important as new standards place emphasis on vocabulary and comprehension of complex texts, as well as discussion as an instructional focus (CCSS, National Governors Association Center for Best Practices Council of Chief State School Officers, 2010). These new standards draw from empirical research that treats academic discussion, a dialogic tool of instruction, as an academic skill in itself as well as an instructional activity to support content understanding and reading comprehension. This observational study pulls together findings from the empirical literature on pedagogical talk moves that have been characterized as supporting dialogic pedagogy, and

shown to relate to language and literacy outcomes. The coding scheme developed from this literature was applied to observational data, and rates of talk moves were parsimoniously tested in multilevel models to predict student reading comprehension scores. Statistically significant talk moves were then explored to describe their qualities and how they function in the pedagogical activities of upper elementary language arts classes.

The overarching hypothesis guiding this study was that high rates of dialogic instructional moves used during instructional talk predict student reading comprehension. This hypothesis was only partially born out by the models and subsequent qualitative analyses. Statistically significant relationships between five talk moves and student reading achievement were found in the final models, but these results provide somewhat conflicting evidence on the relationship between dialogic talk moves and student reading achievement. Using a convergent mixed method analysis (QUANT → qual) allowed for more depth of exploration into these findings. Taken together, these findings are discussed in this chapter.

This chapter is organized into eight sections. The first briefly reviews the general patterns of instruction in these classrooms that framed and mitigated the function and rates of talk moves observed during instruction. The next sections present three explanations for the results of the mixed analysis. The chapter concludes by addressing the implications for teaching, the limitations and future directions of this research, and a final summary answering the research questions guiding this study.

Overarching Patterns in Language Arts Lessons

The importance of dialogic instruction for reading, from the perspective of the simple view of reading, is that it can foster linguistic comprehension skills. This relationship rests on a sociocultural theory of language learning, where interaction and engagement with others

facilitates language development (Vygotsky, 1986), which, in turn supports literacy skill (Chang-Wells & Wells, 1993; Nystrand, 2006; Wells, 1990). Research into second language acquisition explains this theoretical position by emphasizing input (Krashen & Terrell, 1983; Long, 1981) and output (Swain, 2005) for linguistic competency to develop. Thus, to develop linguistic comprehension skills to support reading skill, the learner should have rich opportunities for linguistic input and output, which necessarily means interaction between the learner and others. This perspective on the acquisition of linguistic comprehension skill was only partially supported by the findings in the current study.

Instructional interactions were framed at a 'global' and an 'utterance' level of analysis. Certain patterns of interactions were commonly used in these lessons and tended to structure the way pedagogical talk was carried out at the utterance level of analysis. Using this global lens to analyze interactions, a first finding was that the vast majority of the lessons observed were teacher-managed learning activities accomplished most frequently through a question and answer talk pattern, grounding much of the instruction in these upper elementary classrooms in an IRE model of instruction (Cazden, 1988; Mehan, 1979). This model of talk has its drawbacks and critiques, including its fast-paced questioning (Cazden, 1988), and its emphasis on recitation of content rather than allowing students more control over turn-taking and text interpretation (Chinn et al., 2001). From a dialogic point of view, this instructional pattern inhibits opportunities for student language output that could support their linguistic comprehension growth.

For instance, academic discussions were rare in these classrooms. Discussion is an instructional activity that indicates an emphasis on dialogic instruction. Academic discussions have been lauded in the literature for their disruption of traditional patterns of discourse, like IRE, and their resulting increase in student talk (Chinn et al., 2001), as well as their correlation

with higher content and literacy acquisition (Applebee et al., 2003; Murphy et al., 2009). Definitions of academic discussion in the literature vary, but they have in common a free exchange of information among students and/or teachers (Nystrand et al., 1997), characterized by exploratory talk (Elizabeth et al., 2012) and understanding created by the group as the teacher gives up some control over the content and form of discussion (Billings & Fitzgerald, 2002). Outside of intervention studies, how discussions get started is not fully understood, though some research suggests that authentic questions and uptake can accomplish this in secondary contexts (Nystrand et al., 2003). Like much naturalistic research in classrooms (Applebee et al., 2003; Nystrand & Gamoran, 1991), the 31 classrooms in the current study showed very few instances of academic discussion, either intended by the instructors, or precipitated by dialogic talk moves. Though many classrooms in the Maryland sample used small reading groups during their language arts time, a time that could be more conducive to text-based discussions, much of the talk during these episodes was to read the text aloud or answer test (recitational) questions asking the students to identify information directly from the text, or questions that asked students to infer meanings from the text. Only one teacher in the sample explicitly used discussion groups during one observation, but these literature circle groups were off-task for much of the lesson. Academic discussions that did occur in these classrooms appeared to be not only infrequent, but of short duration as well. For instance, a few teachers consistently offered their students' opportunities to pair up and share their answers to the teachers' questions during Q & A episodes, but these rarely lasted more than one minute. The few discussions that were more substantive and lengthy had the teacher as a participant. Appendix C includes one example of a text-based discussion of the novel *Stone Fox* (Gardiner, 1980) that includes indicators of dialogic instruction, such as uptake questions, and a few instances of extended student talk. However, on

the whole, text-based discussions were infrequent and this was likely due to the heavy reliance on teacher-managed, Q & A patterns of instruction.

In addition to the ubiquitous use of teacher-managed talk patterns, the use of certain pedagogical materials had an effect on the utterance-level talk moves explored in the HLM models. In fact, pedagogical materials and the fields (topics) of lessons may have been an equally strong influence on the consistency or lack thereof in a teacher's use of certain utterance-level talk moves. For instance, when analyzing the within classroom data, comparing lessons by the same teacher across the academic year, teachers with consistently high-rates of talk used materials that emphasized reading (mainly) narratives, with a pedagogical focus on text comprehension in their language arts lessons. Teachers that exhibited lower rates of the talk moves under consideration in this study showed more diversity of topics addressed and materials used, with a trend toward more form-focused and vocabulary work. In particular, instruction that used these two types of materials was rarely accompanied by uptake and high-quality questions.

Another general pattern across classrooms was that teacher explanations rates were higher when informational texts were the focus of the learning activities. Teachers tended to use more extended talk, which included definitions of concepts without posing questions or addressing student input, when engaged with these texts than with narrative genres. Systemic functional linguistic theory suggests that form and meaning are mutually constituting, and the mitigation of talk through materials provides some evidence that teachers in this sample used text genre to construct types of meaning in these lessons. In SFL theory, language and context are systematically linked (Martin, 1992), suggesting that if informational text genres were accompanied by high-rates of teacher explanations, this indicates that this pattern of language shapes a 'correct answer' type meaning from reading these texts. Teachers tended to treat

reading comprehension as an accuracy and efferent activity through this sort of discursive move. Similarly, form-focused materials (mainly grammar worksheets) rarely included uptake questions or teacher explanations. When viewed from an SFL perspective, language is a system of choices that users deploy to make meaning (Christie, 2002). This includes authors of texts, and their readers. Language in worksheets is decontextualized and focused on form rather than meaning, and it follows that teacher questions to do with this content are necessarily form focused and surface-level. This treatment of this genre of pedagogical materials suggests that meaning constructed around these texts was not the result of reciprocal talk or extended definitions.

Though the heavy use of the IRE pattern of instruction, and the use of some kinds of pedagogical materials, likely reduced the opportunities for academic discussions in these language arts classes, the IRE pattern of instructional talk did allow for some dialogic instructional moves to support student reading comprehension. Uptake questions were frequently used in many classes, predicting higher reading scores on the outcome measure, and were used concurrently with other talk moves that, though not dialogic in nature, had a positive impact on these scores. Explanations for the effectiveness of these talk moves are discussed in the following sections.

Overall, the overarching patterns in this sample of language arts lessons revealed a preponderance of teacher talk and a corresponding lack of academic discussions. These findings suggest an atmosphere of instruction focus on the production and dissemination of teacher language, and less emphasis on a the elicitation of extended student talk.

Language Exposure as Support for Student Reading

This lack of academic discussion, while indicative of a frequent pattern of teacher-managed, recitational instructional discourse, did not obstruct some teacher effects on student reading comprehension achievement. The IRE pattern can be operationalized for a wide variety of learning tasks, including dialogic forms of talk where teachers and students can, to varying degrees, co-construct knowledge (Nassaji & Wells, 2000). For instance, within the IRE discursive pattern, some studies have pointed to the effectiveness of dialogic talk moves like uptake questions (Applebee et al., 2003; Boyd & Rubin, 2006; Nystrand et al., 2003) and high-quality evaluations (Nystrand et al., 2003). In the current study, the uptake move was consistently used within this IRE pattern in many classrooms to correlate with higher reading comprehension achievement. In addition, high rates of teacher explanations also showed a positive relationship with student reading, lending support for a language-based model of reading that emphasizes the importance of linguistic comprehension (Hoover & Gough, 1990).

While the uptake move was defined by its reciprocity, and teacher explanations by their lack of it, they are both sources of language around text comprehension, which suggest the importance of teacher language as a mediator for student linguistic comprehension in these classrooms. The way linguistic comprehension was facilitated in these classes, as exposure to teacher language rather than through extended student talk (and therefore more reciprocal discourse), suggests that these data are somewhat at odds with theories of language development that emphasize reciprocity between interlocuters (Vygotsky, 1986; Wells, 1999) and language output (Swain, 2005). However, this finding may point to important limitations to the study, which are reviewed below. The mixed analyses point to several aspects of instruction in these classrooms providing linguistic exposure to aid students in improving their linguistic

comprehension skills, which in turn support reading comprehension (Hoover & Gough, 1990). The talk moves found significant in the HLM models, including teacher explanations, student explanations, and uptake questions, as well as their functions within instructional discourse, suggest their function as input for student language development.

Teacher explanations.

Teacher explanations were the strongest predictors of reading comprehension in this sample, which lend support to the hypothesis that listening comprehension is a source for linguistic skill for reading comprehension (Hoover & Gough, 1990). Teacher explanations have been found to predict young students' language knowledge, a key support for reading. One common feature of teacher explanations in the present study and others is the decontextualized nature of this talk move, a basis for a language-based theory of literacy development (Snow, 1991). Over the elementary years, Aukrust (2007) found that teacher explanations predicted students' later vocabulary knowledge. Similarly, her colleagues found that teacher-led talk predicted same year vocabulary differences in L2 learners at age five, and these differences held at age 10 (Rydland et al., 2013). Other research on young learners has shown that the use complex vocabulary and semantic structures predicts linguistic comprehension (Dickinson & Porche, 2011; Huttenlocher et al., 2002). In the present study, definitional vocabulary instruction was one consistent focus of teacher explanations, a finding consistent with other observational research on vocabulary instruction (Blachowicz, 1987; Scott, Jamieson-Noel, & Asselin, 2003; Silverman et al., 2013; Watts, 1995). Providing vocabulary definitions was not directly related to reading comprehension in this study, but while this study does not map a direct relationship to reading, the positive relationship between teacher explanations and students reading scores is suggestive of the importance of vocabulary instruction for improved vocabulary knowledge

(Proctor et al., 2011; Silverman et al., 2013), which in turn has been extensively linked to reading comprehension outcomes (e.g. Dickinson & Porche, 2011; Nation & Snowling, 2004; Ouellette, 2006; Proctor et al., 2012).

In addition to providing exposure to academic language, teacher explanations were responsive to students' perceived needs. These moves were non-dialogic but responsive to student conceptual knowledge and vocabulary comprehension. By explaining or summarizing a concept, or modeling a strategy, teachers were extending a topic (a text-related vocabulary word for instance) and providing the language learning needed in reading comprehension. Support for this conclusion of the function of teacher explanations for reading is not only from the qualitative analyses identifying the content of explanations, but also provided by the final baseline model of Level-1 predictors in the current study. This model (5; Table 4.10) shows the importance of word recognition, semantic, and syntactic knowledge for reading comprehension. Controlling for students' initial reading comprehension scores, this model shows that semantic-syntactic knowledge and word reading fluency explain 48% of the variance in reading scores, when accounting for students clustered in classes. This provides further support for the simple view of reading; that linguistic comprehension (in the form of semantic and syntactic facility) and decoding/word recognition account for reading differences in this sample of students. From this perspective, the language provided by teachers who used high rates of explanations indicates, through both quantitative and qualitative perspectives, its importance for student outcomes by providing exposure to academic language and language arts concepts.

Direct Instruction.

Teacher explanations had the strongest predictive relationship with students' reading scores, and these explanations mainly provided students with definitions of vocabulary and

descriptions of concepts in a given text, and to a lesser extent, functioned to review previous content, provide "meta" comments about class work, and model reading strategies. The thread running through this talk move was the use of teacher language to directly and explicitly provide content support. Though these instances of explicit, univocal address to students does not meet Duffy and Roehler's early definition of direct instruction as having "an academic focus, precise sequencing of content, high pupil engagement, careful teacher monitoring and specific corrective feedback to students" (Duffy & Roehler, 1982, p. 35), teacher explanations in this sample demonstrated the academic focus, monitoring, feedback, and teacher control of the learning situation (Baumann, 1983) involved in this discursive move. The quality and rate of this sort of direct teaching seemed to play an important role in providing linguistic and conceptual input for students when engaged in language arts activities (in which text comprehension and vocabulary instruction were the majority of the lessons' topics). The literature on direct instruction and reading comprehension has focused on explicit instruction of reading strategies to ameliorate the traditional tendency of teachers to ask assessment questions rather than "create an understanding of either the main ideas or strategy for figuring out the main idea" of a text (Gersten & Carnine, 1986, p. 71). While this was rarely the content of teacher explanations in this sample of classroom observations, the explicit provision of content follows the same instructional genre: students are provided with content (and language) rather than generate it themselves.

During the 1970s and 1980s, direct teaching of reading was lauded as a key support for comprehension, which was extended to include more jointly determined "transactions" around texts (Pressley et al., 1992), and later called into question for its focus on strategies over the actual meaning of the text (Pressley, 2000). However, this research has provided evidence that direct instruction of reading was an impactful support on students' identification of main ideas

(Baumann, 1984; Stevens et al., 1991), and reading outcomes when part of a culturally responsive instructional program (Tharp, 1982). On the other hand, current research on kindergarten reading instruction using multilevel modeling has found that teacher demonstrations of reading skills were not related to early reading skills (Smolkowski & Gunn, 2012), though in this study and an earlier paper, the authors did find that the rate of practice in word recognition, reading fluency, vocabulary, and comprehension skills did relate to student reading outcomes (Gunn et al., 2005; Smolkowski & Gunn, 2012). Furthermore, the lack of a relationship between teacher demonstrations and student reading outcomes in the later study may have been mitigated by low reliability scores and the limited measurement of these talk moves as rate and not quality (Smolkowski & Gunn, 2012).

So, while direct instruction may be a somewhat dated focus of reading comprehension research, it cannot be discounted for its established connections with reading comprehension outcomes. The present study provides some corroborating evidence of this relationship, and direct instruction is conceptualized here as effective not for fostering reading strategy use (though this was the focus of a small number of explanations in this sample of classrooms), but for its attention to students' comprehension of text through the provision of vocabulary definitions and their accompanying models of academic language for the development of student linguistic comprehension.

Student explanations.

Other evidence for the relationship between students' exposure to teacher language and their reading comprehension outcomes was the finding that extended student talk predicted lower reading scores. The vast majority of the student explanations met only the minimum requirements set for applying this code to the transcript data, which were rather brief two- to

three-sentence utterances usually focused on answering a teacher's question about a text. Student explanations as a feature of instructional discourse may be negatively related to student reading achievement because of the brevity of utterances that limit students opportunities to use language, as language output is theorized to be a second component important to language development (Swain, 2005). Thus, in light of the previous discussion on teacher explanations, and following the simple view of reading, it would seem that language exposure through teacher input was a more important effect on student linguistic comprehension development in these classrooms. Why opportunities for linguistic output would be negatively correlated with reading outcomes when output is a key part of language acquisition theory may have less to do with disconfirming the theory and more to do with the highly correlated nature of the talk moves (see Table 4.9). This limitation is addressed in a following section. However, it should be noted that nn addition to the limited quantity of student talk, accountability for student clarity and descriptiveness during was not always apparent during student explanations; many of these explanations were personal experiences shared with the class but not pursued by teachers or peers to provide opportunities for students to extend their knowledge base through this sort of linguistic output.

Uptake questions and inference.

In addition to the talk moves of uptake questions and teacher explanations as sources of language input to benefit student linguistic development, uptake questions promoted the skill of inferencing necessary for reading comprehension, and a particular skill required by the WMLS-PC outcome measure in this study. The high number of uptake questions pressing students for inferences were a leverage point for students to develop linguistic comprehension, since successful inferencing skill has been show to rest on vocabulary knowledge (Cain et al., 2004).

Vocabulary exposure has been consistently positively correlated with vocabulary breadth (e.g. Hart & Risley, 2003; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991), which in turn is a strong predictor of reading comprehension. Thus, it is likely that vocabulary knowledge is supporting reading comprehension through both the direct teaching of definitions, as well as through questions pressing students to apply their vocabulary comprehension to infer information from texts. Uptake questions and explanations may be two talk moves working in tandem to provide students with consistent exposure to academic language and practice using a reading strategy demanded on the outcome measure.

This indirect pathway to reading comprehension is one explanation for why the frequency and quality of uptake questions, and the frequency of teacher explanations were strong predictors of student reading achievement. Talk moves such as uptake questions provided students with both linguistic exposure and opportunities for the application of their vocabulary comprehension as well as practice with a key reading strategy afforded by the focus of this kind of teacher question. The focus on inferential comprehension through uptake in these classes was consistently applied to narrative texts. When inference questions were not scaffolded to test-like questions (as in the case of many high-quality questions), they related positively to student reading achievement, corroborating other findings in instructional effects on reading. Drawing on the same sample of classrooms but using a different coding scheme, Silverman and colleagues (2013) found that instruction asking or guiding students to make inferences was positively related to growth in reading comprehension. This analysis and the present analysis align with reading interventions that instruct students on this crucial skill (e.g. Beck et al., 1996). However, this finding of the possible importance of inference questions during instruction in the current study is tentative. When HLM models were run with only uptake or high-quality questions as a

single level-two predictor, there was no significant relationship with reading. This suggests that these questions may be capturing the effects of other talk moves or content.

In addition, there was some evidence the data to suggest that the efficacy of these moves, and hence language exposure for students, may be mitigated by the pedagogical materials used during these lessons. Teachers rarely posed uptake questions when working with form-focused and vocabulary worksheets, suggesting that a focus on the comprehension of narrative texts and test-prep materials benefited from some reciprocity between teacher and student interactions. It may be that highly structured and decontextualized texts may limit opportunities for reciprocity of talk, as suggested by SFL theory that form and meaning are mutually constituting, but there is little evidence of this in the empirical literature.

High-quality questions.

A corroborating finding for evidence of language exposure as a key support for student reading comprehension was a significant predictor of lower scores. High-quality questions were associated with low reading scores in the final models, disconfirming the hypothesis that this dialogic move would predict higher scores. While most of these high-quality questions asked students to infer meaning from texts like their uptake counterparts, which would seem to benefit reading comprehension (see Cain et al., 2004), these questions indicated a pattern of scaling down their possible positive impact through scaffolds. Qualitative analyses of these questions indicated that many teachers posed these questions and quickly followed them with lower quality questions that asked students to recall or recite (rather than infer or analyze, for example) or answer more narrowly-focused questions, resulting in students' contributions that were more attenuated than if answering the original high-quality question. However, while scaling down a question means more teacher talk overall, which would corroborate the working theory here that

significant talk moves were acting to provide students with linguistic exposure to aid comprehension, teachers were reducing the content exposure for students. There is some suggestion in the literature that too much scaffolding may be counterproductive for the development of linguistic skill (McElhone, 2012). If this is indeed how high-quality questions were functioning in these data, it is probable that teachers were adjusting their instruction to meet their students' perceived language abilities, a typical scaffold of adults with children in general (Weizman & Snow, 2001). However, these adjustments were not providing students with opportunities for some higher-level thinking like inferencing, generalizing, or analysis.

To further explore this perplexing finding, high-quality questions were run as single predictors in an alternative multilevel model and found not to be significant, suggesting that when in the final model with other talk moves, these questions were likely capturing other facets of talk or instruction that were lowering students' reading scores on the measure. High-quality questions were highly correlated with uptake questions ($r = .80$, $p < .001$), resulting in a possible suppression effect. Future analyses must interpret the two coefficients together because they were so highly correlated. Since there is no established theory for an order of entry for these classroom level variables when modeling these talk moves, we entered them at once, which likely resulted in oddities such as high-quality questions predicting lower reading scores. A next step will be to recode high-quality questions in absence of uptake and high-quality questions in the presence of uptake to explore this relationship more carefully.

Summary.

Exposure to academic language for linguistic development is well tread empirical ground, and this study provides further evidence of this. While language tasks like academic discussions and high rates of interaction between interlocuters may be ideal sources of language

development, language exposure limited in reciprocity may still have positive effects on student reading achievement.

Reciprocal (uptake questions) and monologic talk moves (explanations) suggest some complexity in categorizing instructional talk moves. Both afforded students with exposure to language around texts and practice with reading strategies typically seen on standardized reading measures like the WMLS-PC. Teacher explanations, in particular, attended to students' academic vocabulary and conceptual knowledge, providing models of language that fostered linguistic skill and vocabulary breadth necessary for reading comprehension. Uptake questions provided student with practice inferencing information from (mainly) narrative texts, as well as reinforced the importance of vocabulary knowledge. This knowledge plays an essential role in reading comprehension outcomes through elementary school (Nation & Snowling, 2004; Ouellette, 2006; Proctor et al., 2012). Though more analyses are on-going, high-quality questions may also support this view that significant talk moves acted as a source for students' linguistic comprehension, as these questions tended to dilute the strategy of inferencing toward recitation of known information, which did not allow students to engage with this reading strategy while providing little in the way of extra language exposure around texts.

More analytical work is needed to conceptualize how these talk moves may work in tandem to provide students with linguistic skills that are leveraged into their reading comprehension tasks. However, at this point in the analyses of these data, students' linguistic comprehension seemed to be facilitated through exposure to teacher language more so than through extended student talk, as both uptake and teacher explanations shared attributes of providing sources of linguistic exposure that were opportunities for students to develop linguistic comprehension, a key part of reading comprehension (Hoover & Gough, 1990).

Teacher Talk for Maintaining Attention to Literacy Tasks

The importance of language exposure for reading comprehension during literacy instruction may be aided by talk moves that encourage attention and accountability to the learning tasks. There is some suggestion that techniques of teachers to attract and hold students attention are related to comprehension (Dickinson & Porche, 2011). The rates of uptake questions and teacher evaluations function to affirm student contributions, talk moves which are reasonable candidates to attract and hold student attention, bringing students into the lesson discourse. In addition, the quality of student explanations in this sample also supports the notion that attention to others' contributions about texts may be a scaffold for young readers.

The significance of uptake questions in this study corroborates other studies, mainly in secondary classrooms, that have revealed their importance in supporting literacy tasks (Applebee et al., 2003; Boyd & Rubin, 2006; McNeill & Pimentel, 2010; Nystrand et al., 2003). Their function to press students for explanations, to press for more information, or to solicit more participation may be a factor in reading achievement as a talk move to maintain student attention to literacy tasks.

This possible function aligns with the two other talk moves found to be predictive of positive scores. The affirmative quality of low-quality teacher evaluations demonstrated teacher responsiveness, albeit brief, to what contributions student were making. Positive reinforcement was one feature of literacy instruction that has been found effective in first grade classrooms (Pressley et al., 2001). And teacher responsiveness has been linked to teacher supportiveness for language development (Dickinson & Porche, 2011). In a longitudinal study of language development from preschool to grade four, Dickinson and Porche (2011) found that teachers' efforts to maintain student attention on learning tasks directly related to later reading

comprehension. In the current study, while low-quality teacher evaluations in and of themselves did not attend to the content of the lessons in substantive ways, the majority of them were immediately followed by questions that furthered the lesson along, keeping the lesson focused on the text. It may be that talk moves such as uptake and evaluations, as affirmations of student contributions, worked to maintain student attention and keep the lessons moving, and provided direct support for student outcomes during the same instructional year. Though there was little language or content used in these low-quality evaluative moves, lessons presumably would not move along without these acknowledgements of student contributions, particularly as so many instructional episodes followed the classic IRE pattern of instructional talk. Student contributions were evaluated before the lesson could move on, and in this sense, provided students with a judgment or an acceptance of the appropriateness of their answer, and then moved to direct students' attention elsewhere along this IRE, question and answer pattern of lessons.

A second finding of the value of instructional talk to maintain student attention on language arts tasks was corroborated by the significance and features of student explanations. Extended student talk was hypothesized to be an indicator of dialogic practice during instruction, but was found to predict lower reading scores. This finding was surprising until patterns in the data showed that these explanations rarely met more than the minimum requirement for the code (two lines of transcript or two ideas), and rarely followed moves like uptake (only 2% of uptake questions were followed by student explanations), that frequently pressed for explanations and more information. This output from students contained few instances of exploratory or analytic talk, lauded as publically accountable, constructively engaged and cooperative talk where interlocuters share their reasoning (Elizabeth et al., 2012), but did respond to teacher questions about content of texts. Student explanations functioned more like cumulative talk, where students

may build briefly on others' ideas but not engage with analysis of a topic (Elizabeth et al., 2012; Mercer et al., 1999). For instance, the content of student explanations was frequently the sharing of personal experiences, which were not always on topic, and were organized to have as many students share their stories as a string of loosely connected vignettes. Such a use of extended student talk could be an indication of the teacher not maintaining students' attention to the topic of the text or lesson (Hoff-Ginsberg, 1991). It is thus plausible that a lack of exploratory or more in-depth talk on a subject and a lack of building onto student explanations may have distracted or disengaged students from the learning focus, as students in classrooms with high rates of this move showed lower reading comprehension scores.

However, a conflicting finding from the qualitative analyses linked low-quality evaluations, argued above to aid students' focus on the lesson, with student explanations. Classrooms with high rates of low-quality evaluations had about 100 more instances of extended student talk (student explanations) than low-rate classrooms. In these classrooms, it may be that the quantity of evaluations trumped the amount of student explanations because of the lack of quality and focus on texts in the explanations. It has been suggested in the literature that the occurrence of low-quality teacher evaluations may be indicative of student talk and instances of discussion (Nystrand et al., 2003), but while there was a pattern of quantitatively more talk when more low-quality evaluations were used, this did not precipitate text-based discussions where analytic, exploratory, or reciprocal talk was evident.

Finally, teacher explanations, while shown to be the strongest predictors of higher students reading outcomes, were more opaque in their function to maintain student attention. In the qualitative analyses, there were no discernible patterns in the discourse that showed explanations resulting from student questions about vocabulary or content, or precipitated by

incorrect answers that needed adjustments or clarifications; high-quality teacher evaluations captured these sorts responsive of teacher moves, but they did not approach significance in the final models. The content of teacher explanations also did not show any pattern of attention-getting moves, like explaining vocabulary with humorous scenarios, or regularly using students' background experiences to explain a concept. While many teachers showed evidence of this sort of content, these teachers did not necessarily show higher rates of explanation that would have predicted their students' higher reading scores.

The Construction and Measurement of Reading

A third explanation for the quantitative and qualitative findings, in addition to teacher talk being a significant source of language exposure and attention management, was the similar treatment of reading in instruction and on the reading comprehension measure. The frequent emphasis of the cognitive skill of inference around text-driven topics mirrored one of the skills necessary for successfully completing the reading task on the outcome measure. The alignment of instructional and measurement tasks may help explain why certain talk moves showed a strong relationship with the standardized measure of reading used in this study, while others hypothesized to predict reading, did not. Correlational studies of specific talk moves have used curriculum-aligned outcome measures of reading comprehension, which have shown a greater number of dialogic moves predicting reading achievement (Applebee et al., 2003). The Woodcock Muñoz Language Scales, Revised, Passage Comprehension (WMLS-PC; Woodcock et al., 2005) measure, and similar standardized reading tests used by both states in this sample, may be limiting the value of dialogic talk moves or more dialogue and discussion in these classrooms.

Significant talk moves and case study analyses of these moves indicate an overwhelming preference of an efferent stance (Rosenblatt, 1994) toward text and reading in this sample. Even questions about aesthetic devices like similes were deconstructed to check for student understanding of the concept, and didn't also respond to such literary devices aesthetically:

> Teacher 10: Can someone tell me a simile (in the reading)?
> Student: A thousand starfish.
> Teacher: Okay, go ahead, say it.
> Student: A thousand starfish in the sky.
> Teacher: Okay. Can you read the whole sentence, Mary? She was, starting with she was? Everyone put your finger on she was.
> Student: She was looking up through the water at the sun sparkling like a thousand starfish in the sky.
> Teacher: Very good, you were listening during simile, comparing two things using like or as. What's she—what's the author comparing in that sentence, Gianni? What are the two things? She was looking up through the water at the sun sparkling like a thousand starfish. What are the two—what's sparkling like a thousand starfish?
> Student: The sun. (P6:131, 1/31/11)

This typical example of an efferent stance toward reading, which was clear in all 31 classrooms, was supported by evidence from four of the five significant talk moves, and concurred with other research that found discussion-based reading interventions in which teachers control more of the activity are efferent in stance (Murphy et al., 2009). Uptake questions focused on students' text comprehension, pressing students for explanations to verify their understanding of a plot event or character's behavior. No questions of any type elicited student evaluations of or thoughts on writing style or their own experiences while reading the text. Teacher explanations affirmed the accuracy of text comprehension by providing students with definitions of key vocabulary and concepts. There were no incidences of teachers modeling with a view to aesthetically experience a given text, for example. Low-quality evaluations most frequently affirmed the accuracy of text and vocabulary comprehension. Conversely, high-quality questions tended to scaffold 'down' inferential questions to have students provide

answers explicitly stated in texts, which clearly established a view of reading as a process of the acquisition of information (Rosenblatt, 1994). The skill of inference was key to successfully scoring on the WMLS-PC, and this function of high-quality questions worked against this, as indicated by its negative relationship with reading comprehension scores across all grades and students.

The WMLS-PC is constructed from the same understanding of reading: students read short passages to complete increasingly more abstract and decontextualized texts by filling in a blank with a semantically- and grammatically-appropriate word. The task emphasizes the need for students to read for extracting information, and has been critiqued for its attention to lower cognitive skills (Francis et al., 2006). It is also strongly related to word reading ability (Francis et al., 2006), which in these grades was not a topic of instruction. Students in this study scored at the age-equivalent standardized score mean on the decoding and fluency measure, suggesting that teachers may have seen little need for addressing decoding and fluency skill during their language arts lessons. Instead, lessons were predominantly focused on text comprehension, and instructional moves emphasized language exposure as a scaffold for supporting this comprehension.

To complete the WMLS-PC tasks, students must draw on their own vocabulary knowledge to complete sentences. Indeed, vocabulary breadth has been shown to be a predictor of achievement on the WMLS PC (Montecillo Leider, Proctor, Silverman, & Harring, 2013). As argued above, the large amount of academic teacher talk that students were exposed to in these classes likely aided their vocabulary knowledge. The measure may also have a negative relationship with the skill of narrative production, where the reader integrates information from schema into narrative forms for retention (Francis et al., 2006, p. 319). Instruction in these

classes rarely exhibited a focus on students producing narratives, either orally or in writing, as there were low rates of extended student talk across the sample. This measure thus taps decoding and vocabulary skills, two skills where students in this sample test close to the standardized sample mean. In the case of the latter skill, pedagogical discourse shows a high frequency of language exposure from teacher talk, a source of vocabulary learning (Beck, McKeown, & McCaslin, 1983; Silverman & Crandell, 2010).

Thus, the emphasis on language exposure through teacher talk, more so than a reciprocity between teacher and student, and the alignment of the most frequently used strategy attending to textual meaning (inferences) reflect how reading comprehension was operationalized in this study. This view of reading in instruction and on the WMLS-PC likely influenced how the relationship between talk moves and reading outcomes functioned. If the reading measure was only capturing a narrow range of reading skill, it may be limiting the influence of dialogic talk moves like high-quality evaluations on reading scores.

Models of reading comprehension and reading instruction have developed along similar theoretical trajectories: both have shown the early influence of cognitive, skill-based processing models, and lately have turned toward more situated, socially-constructed models of reading, though perhaps models of reading comprehension have been a bit slow to catch up. As it stands, research, such as the present study, must make do with reading comprehension instruments that tend toward efferent stances of text. One would hope however, that dialogic instruction and dialogic talk moves such as those explored in this research, support students to not only achieve on these tests, but to move beyond this treatment of reading comprehension toward critical, evaluative and aesthetic stances toward texts.

Limitations to the Study

The findings reported in this study include a number of limitations. The study is the first to pull together pedagogical talk moves (stemming from a variety of research methodologies) into one coding scheme, and thus is more exploratory in nature and its findings tentative. These talk moves have not been modeled in this comprehensive way before, nor at these grade levels. In addition, the codes applied to these data were derived from research on dialogic instruction, but the data showed little dialogic and reciprocal events; students did not talk to each other about texts or language arts topics, and teachers did not conduct sustained, reciprocal text-based discussions either. As such, the mixed results confirming and disconfirming our hypotheses may be a result of analyzing phenomena that were not in these data in any substantive ways.

The quantitative analyses used here rest on the assumption that there are direct pathways from instructional talk to outcomes, and while this assumption may help to identify talk moves that relate to student outcomes like standardized reading measures, talk moves exist within the larger contexts of teacher epistemology (Nystrand, 2006) and resulting choices about materials, topics, and learning tasks that may mitigate instructional talk moves. For instance, instruction could have been influenced by teachers' perceptions of student comprehension levels (Silverman et al., 2013). Establishing closer links between discourse and reading skill would be possible with careful discourse analysis tracking individual students across their lessons, but lose the power of a larger sample size. Correlational analyses do not account for unmeasured variables, and while there are many possibilities of mitigating variables at work in these classrooms, this is one reason why qualitative analyses were conducted. Further modeling of such pedagogical resources on these dialogic talk moves can aid our understanding of factors influencing both talk and outcomes.

A third limitation of these analyses is the coding scheme itself. Despite good reliability on nine of the original eleven codes, these talk moves indicated some overlap in their definitions and use. Due to the highly correlated nature of the Level-2 predictors, these individual findings may be less to do with the individual talk move itself, and more about the talk move working in conjunction with other talk moves to predict student achievement. To test this, models were run that only included each of the nine talk moves individually, and only low-quality questions and low-quality evaluations were significant in these HLM models. These were the more frequently applied codes in the data, which suggests the possibility that the model is capturing something about the quantity of instruction, rather than the quality of instruction implied with these codes, which may corroborating with other findings (Smolkowski & Gunn, 2012). There may have been a lack of power for some individual talk moves by virtue of the limited use across classrooms. This could suggest a coding scheme that may be characterizing instruction too parsimoniously. A second result of the highly correlated nature of talk moves may have been suppression effects. A more appropriate way to conceptualize many of these moves may be on a continuum rather than a dichotomous scale (Elizabeth et al., 2012). On the other hand, lower-frequency codes like high-quality questions and explanations became predictors in conjunction with other talk moves, which suggests further analyses that include interactions between Level-2 predictors. The qualitative analyses also support the assessment that interactions between moves may capture more about the indirect factors influencing reading outcomes. Modeling interaction terms is a future step with this dataset.

A corollary to the limitation of this conceptualization of the codes is the use of rate in the quantitative models. Using the rate of a talk move (amount of a given talk move per minute of instruction) ignores the fact that some students are exposed to more of all kinds of talk than other

students. It may be that dosage, rather than rate, is more of a factor in scaffolding student reading comprehension, though well-regarded research on instruction and literacy has used frequency-based variables in HLM models (Applebee et al., 2003; Smolkowski & Gunn, 2012).

Another limitation is that the classrooms observed were a sample of convenience, chosen from a pool of willing teachers that had a number of students participating in the parent CLAVES study. For generalizability, these classrooms should have been randomly sampled. It is important to note that a comparison of the observed and non-observed classrooms in the larger CLAVES sample showed that on some measures, the observed classes were statistically significantly lower than the non-observed classrooms. This may have been due to the sampling of students at both sites. Since both the larger study and the present study were interested in bilingual students, we purposively oversampled bilingual students, including those classified as English learners. It is well established that this subpopulation of K-12 students in the US shows much more variability on student literacy outcomes than their monolingual counterparts and this increased variability may have affected the classroom mean scores of the observed classrooms. While this is an important limitation for larger studies aiming to generalize to the US student population, this difference in classrooms is not of such concern for this small, observational study.

A final limitation is the use of a Kappa statistic for establishing inter-coder reliability on the coding scheme. Hill and colleagues (2012) note that when examining rater quality through the measurement of inter-rater reliability on classroom observation data, inter-rater reliability statistics may mask other sources of variation in observational data like the kinds of lessons that can affect observations of teacher behaviors. A single inter-rater reliability statistic also "fails to estimate interactions between raters, teachers, and lessons" and thus may not provide a more

comprehensive picture of score reliability (p. 57). With this in mind, we calculated reliability statistics for each talk move for more precision, but since the focus of the study was on talk moves, we did not score reliability on other variables like lessons and teachers. A further limitation is that the Kappa statistic is influenced by the number of scored items; some items can have high agreement with few items in total, and vice versa. In spite of this possible limitation, reliability established in this study followed the norms of inter-rater reliability calculation in similar research studies, and was chosen because of its well-established use in this sort of empirical research.

Implications for Instruction and Future Directions for Research

While these are exploratory and site-specific analyses that are not readily generalizable, these analyses do suggest some implications for the support of upper elementary students' reading comprehension. Most broadly, the teacher has an important role in providing students with academic language exposure. Through the explanation of content, the choice of pedagogical materials that frame instructional talk, and the management of instructional discourse, teachers can provide rich or poor contexts for the development of students' linguistic comprehension.

Secondly, these teachers used their talk to provide linguistic exposure in a climate of attentiveness to student responses to maintain students' own attention to learning tasks. Through immediate feedback, even with low-quality evaluations, teachers can maintain student attention on pedagogical talk by affirming their participation and contributions. Talk moves like uptake questions and high-quality questions hold the potential to meet both the need for language exposure and on-task attention for successful reading comprehension. For instance, uptake questions can press students for inferences while providing students with a model of a key reading strategy and text-based, content-focused language.

Instructional talk patterns across these classrooms regularly followed the oft-critiqued discourse pattern of IRE (Cazden, 1988; Mehan, 1979), concurring with research in other settings (e.g. Applebee et al., 2003; Nystrand & Gamoran, 1991; Nystrand et al., 1997). However, within this ubiquitous talk genre, there is potential for reciprocal and dialogic interactions between teachers and students. Uptake questions and similar moves in the 'E' position, that press students for more talk and higher levels of thinking, may be the next step to move traditionally organized instructional talk toward reciprocal and discussion-based exchanges between actors in the classroom, as suggested by other research in this area (McElhone, 2012; Nassaji & Wells, 2000). These findings emphasize the value of teacher training that includes a focus on pedagogical language, and suggests the importance for individual teachers to incorporate into their lesson planning follow-up moves and other talk moves that take up student contributions and provide academic language exposure.

Questions that elicit student high-quality cognitive skills like inferencing and synthesizing may also be worth advanced planning, though a drawback of this is a research focus on ameliorating teacher questions because of the difficulty in reliably categorizing and coding question types. A different approach to analyzing instructional effects on literacy development may be to look at what students do and then see what teachers are doing to encourage this. For instance, while higher rates of student explanations predicted lower reading comprehension scores in this study, it would be unreasonable to disregard the importance of student talk at these grade levels. Instead, we can look at what encourages student talk on task, and analyze students' own responsiveness, and not just the teachers'. After all, it is what *students* do that ultimately matters for reading research. For instance, in posing high-quality questions, teachers expect students to generalize or synthesize. However, if students do not follow these questions with the

use of these higher-level cognitive skills, the question does not matter. An added variable to this may be the need to better understand developmentally-appropriate instructional talk (Lawrence & Snow, 2011). If researchers can describe what students tend to say and do at all grade levels, we may find that students may not be developmentally ready to talk with each other in true reciprocal fashion at certain ages.

Another implication for teaching and further research is the necessity of improving reading comprehension measures to capture higher level cognitive skill like syntheses or analyses, or to capture how readers track main ideas. If linguistic comprehension is a key component of reading comprehension, as it is argued in this study, capturing how readers track their comprehension based on their language skills would be a valuable source of guidance for reading instruction and instructional talk.

Finally, with regard to future directions of the current dataset, site differences, interactions between talk moves, and the possible mitigating effects on talk from pedagogical materials are three threads to follow from this exploratory study. When comparing high- and low- rate classrooms, the Maryland (MD) teachers were always more frequently represented in the low-rate classrooms on all five significant predictors of reading. Clearly, some site differences are at work in these findings. It seems that MD teachers were not using as many talk moves, or as much of a variety of talk moves, as the Massachusetts (MA) sample teachers. This may have been due to the MD district literacy curriculum that many teachers in the sample followed. Instruction during these lessons seemed highly structured and formulaic, so that some classrooms looked very similar in terms of talk and talk patterns. This was mainly around vocabulary instruction. It could be that this curriculum is restricting teachers' opportunities to use higher rates of talk moves, or a larger variety of talk moves compared to the MA sample,

where teachers seemed to have more instructional freedom. However, a recent meta analysis of reading programs in the upper elementary grades suggested that the specific curriculum program used in the MD district had positive effects on student reading and vocabulary, though this could be due to the extensive professional development provided to teachers in both studies reviewed (Slavin, Lake, Chambers, Cheung, & Davis, 2009). Site differences were beyond the scope of this study but would be worth further exploration to identify institutional factors mitigating the frequency of talk moves in the districts.

Discussed previously, further exploration of each talk move is needed to better understand how certain talk moves may be interacting with one another to produce some of the more counterintuitive results of the multilevel models. Uptake questions and high-quality questions were highly correlated and yet predicted opposite reading outcomes. The highly correlated nature of many of these utterance-level moves must be addressed to understand any interactions or suppression effects in these data.

A final general pattern in these data worth noting here was the possible relationship between pedagogical materials and the use and rate of talk moves. For example, teacher explanations occurred most frequently with informational texts during instructional episodes that attended to students' reading comprehension. Informational texts were not well used in these classrooms, aligning with other research on the most common texts used in elementary instruction (Duke, 2000). If this pattern is strong enough without many instances of informational text use, it would be important to explore how this text genre is treated by teachers and how it influences question type or follow-up moves. A further impetus for this research is the current focus of new Common Core State Standards (CCSS; National Governors Association

Center for Best Practices Council of Chief State School Officers, 2010), emphasizing informational text use.

Summary of Research Findings

There were five research questions framing this study of classroom discourse and reading comprehension outcomes. The overarching research question guiding this convergent mixed method study was answered by four specific questions framing the quantitative and qualitative analyses, which are briefly summarized below.

Overarching RQ: How does the quantity and quality of dialogic instruction affect student reading comprehension in culturally- and linguistically-diverse classrooms?

Most broadly, this study suggests that the quantity and quality of dialogic instruction affects student reading comprehension through five talk moves, only two of which are explicitly reciprocal and dialogic, but which all provide students, in varying ways and degrees, language exposure for linguistic comprehension skills, moves to encourage attention to language arts tasks, and alignment of instructional moves with the reading comprehension measure. Each of the more specific research questions are answered in turn.

RQ 1: What features of dialogic instruction predict reading comprehension scores?

Three teacher talk moves were significant and positive predictors of student reading comprehensions scores. Teacher explanations, which were not hypothesized to predict scores and not defined as dialogic moves to involve students in reciprocal interactions around texts, were the strongest predictors of higher student scores (γ_{07} (28.46); $p < .01$). To put this in context, a half standard deviation increase in the rate of teacher explanations would predict two standard score points higher than the sample's mean on the outcome measure. Uptake questions and low-quality evaluations were also significantly and positively predictive of reading comprehension

scores (γ_{02} (25.13); $p < .05$) and γ_{06} (10.18); $p < .05$, respectively). The finding that uptake questions predicted positive reading outcomes confirmed our hypothesis on the impact of dialogic instructional moves.

High rates of student explanations and high-quality teacher questions were significantly and negatively predictive of reading scores (γ_{08} (-23.73); $p < .01$ and γ_{03} (-16.69); $p < .05$, respectively). These two findings disconfirmed our hypothesis that dialogic moves would relate positively to reading comprehension outcomes.

RQ 1a: Are bilingual and monolingual students differentially impacted by exposure to high levels of dialogic instruction?

The significant talk moves above were main effects across all three grade levels and language status of the students. Future analyses should confirm this finding by running interaction terms in the final models.

RQ 2: What patterns and variations in the significant findings of dialogic instruction from RQ1a are seen in these culturally- and linguistically-diverse classrooms?

Teacher explanations mainly provided students with vocabulary definitions and descriptions of concepts to support understanding of a given text topic. Explanations occurred most frequently with informational texts, which were not frequently used as pedagogical materials across the sample.

Uptake questions were focused most frequently on students' text comprehension, pressing students for explanations and clarifications to verify their understanding of characters and plot events. High-quality uptake questions were almost exclusively inference questions. Uptake questions were rarely used with form-focused materials.

Low-quality evaluations were the most frequently applied code to these data, accounting for 12% of the total codes. The vast majority were affirmations of student input. Questions of differing types followed evaluations about seven percent of the time, and classrooms with high rates of low-quality evaluations indicated a moderate pattern of more instances of extended student talk (student explanations) than low-rate classrooms. There were no patterns of pedagogical materials mitigating the use of this talk move.

Student explanations were infrequent across the sample, and were either responses to teachers' questions around inferencing information from texts, or sharing personal experiences. The quantity and quality of both types of explanations were low, and students were not regularly held accountable to clarify their explanations. Teacher or peers rarely built on these explanations.

Finally, high-quality questions of varying types of questions have no clear patterns of precipitating or following talk moves. They tend to elicit inferences and speculations, but were also frequently scaffolded to recitational questions that required narrow, known-answer information.

RQ2a: What variations of significant features of dialogic instruction are seen within classrooms?

Pedagogical materials and the fields (topics) of lessons may have been the strongest influences on the consistency or lack thereof in a teacher's use of talk moves. Teachers with consistently high-rates of talk used materials that emphasized reading (mainly) narratives, with a pedagogical focus on text comprehension. Classrooms with low-rates of significant talk moves showed more diversity of topics addressed and materials used, with a trend toward more form-focused and vocabulary work, particularly in classrooms with low-rates of uptake and high-quality questions. The rate of teacher explanations rates was higher when informational texts

were the focus of the learning activities, and low when teachers used form-focused materials like grammar worksheets.

Conclusion

New standards of learning outcomes in the US currently emphasize the critical importance of students reading comprehension across a range of texts (CCSS; National Governors Association Center for Best Practices Council of Chief State School Officers, 2010). At the same time, there is an increasing interest in measuring teacher effectiveness, some of which may be misguided attempts to impose rigid structures on teaching, but is nonetheless an important determination for addressing student learning outcomes. Furthermore, US public schools are becoming increasingly diverse: one in five students speaks a language other than English at home, and 10% are classified as English learners that must receive support to access age-appropriate curricula (NCES, 2013). These trends in education today sit against a backdrop of intractable and persistent achievement gaps between language-majority and language-minority students. The confluence of these factors speaks to the urgency for educational research that is conducted in ecologically valid settings and is aimed at producing rigorous and useable research to support teachers and their students.

This is the express intent of the current study. The research was conducted in typical upper elementary, public school settings in working- and middle-class areas that have become increasingly diverse because of economic and immigration patterns in the two school districts. Teachers in these districts reflect national norms; they are mainly White, English-speaking females (Strizek, Pittsonberger, Riordan, Lyter, & Orlofsky, 2006). The challenge for these teachers and their teacher education programs is to teach culturally- and linguistically-diverse students who may have very different life experiences and skills from mainstream, middle-class

students. Teaching reading, in particular, is crucial for teachers to be effective educators, as reading text, along with oral language, underlie and pervade all instruction. Students must learn through language (in verbal and text form) and demonstrate their understanding through language. Therefore, this study is an attempt to address the lack of empirical studies that focus on the potential of dialogic instruction for supporting language skill and reading comprehension.

Findings from these analyses suggest that the language exposure experienced by students in these third-, fourth-, and fifth-grade classrooms through five talk moves had a positive relationship with reading comprehension. Using the simple view of reading as a framework (Hoover & Gough, 1990), these talk moves suggest their importance in developing students' linguistic comprehension, one of two key components of reading comprehension (the other being decoding). However, while the significant talk moves provided students with exposure to academic language and skills, the quality of their use in these lessons also suggest a lack of opportunities for language output to further support students' development of linguistic comprehension. Teachers in this sample were responsible for most of the language production and for providing a breadth of language and content over depth of content and language use, including a lack of argumentation, analytic, or exploratory talk (Chin & Osborne, 2010; Elizabeth et al., 2012). Teachers were also responsible for using talk moves to maintain their students' attention on language arts tasks, which has been tied to later reading outcomes (Dickinson & Porche, 2011).

Overall, the dialogic nature of instruction, here hypothesized to relate to reading comprehension, was not confirmed by these analyses. Only the dialogic move of uptake questions confirmed the potential of dialogic instruction for reading achievement for bilingual and monolingual students alike. Other talk moves pointed to the promise of discourse that

incorporates students' attention and participation. In all, these findings support the simple view of reading as having explanatory power for a language-based theory of reading comprehension, but the power of dialogic instruction to foster linguistic comprehension in these upper elementary classrooms of culturally- and linguistically-diverse students is still underdetermined.

References

Acock, A. (2010). *A gentle introduction to Stata* (3 ed.). College Station, TX: Stata Press.

Anderson, R., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Anderson, R., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., et al. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction, 19*(1), 1-46.

Anglin, J., Miller, G., & Wakefield, P. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58*(10).

Applebee, A., Langer, J., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40*(3), 685-730.

August, D., Kenyon, D., Malabonga, V., Louguit, M., Caglarcan, S., & Carlo, M. (2001). Extract the Base Test—English. Washington, DC: Center for Applied Linguistics.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Mahwah, NJ: Lawrence Erlbaum.

Aukrust, V. (2007). Young children acquiring second language vocabulary in preschool group-time: Does amount, diversity, and discourse complexity of teacher talk matter? *Journal of Research in Childhood Education, 22*(1), 17-37.

Barnes, D., & Todd, F. (1977). *Communicating and learning in small groups*. London, UK: Routledge & Kegan Paul.

Baumann, J. (1983). A generic comprehension instructional strategy. *Reading World, 22*, 284-294.

Baumann, J. (1984). The effectiveness of a direct instruction paradigm for teaching main idea comprehension. *Reading Research Quarterly, 20*(1), 93-115.

Beck, I., McKeown, M., & McCaslin, E. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal, 83*(3), 177-181.

Beck, I., McKeown, M., Sandora, C., Kucan, L., & Worthy, J. (1996). Questioning the author: A yearlong classroom implementation to engage students with text. *The Elementary School Journal, 96*(4), 385-414.

Bellack, A., Kliebard, H., Hyman, R., & Smith, F. (1966). *The language of the classroom*. New York: Teachers College Press.

Billings, L., & Fitzgerald, J. (2002). Dialogic discussion and the Paideia seminar. *American Educational Research Journal, 39*(4), 907-941.

Birks, M., Chapman, Y., & Francis, K. (2008). Memoing in qualitative research: Probing data and processes. *Journal of Research in Nursing, 13*(1), 68-75.

Blachowicz, C. L. (1987). Vocabulary Instruction: What Goes on in the Classroom? *Reading Teacher, 41*(2), 132-137.

Bloome, D., Carter, S., Christian, B., Otto, S., & Shuart-Faris, N. (2005). *Discourse analysis & the study of classroom language & literacy events: A microethnographic perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Boyd, M., & Rubin, D. (2006). How contingent questioning promotes extended student talk: A function of display questions. *Journal of Literacy Research, 38*(2), 141-169.

Bransford, J., Brown, A., & Cockling, R. (2000). *How people learn: Brain, mind, experience, and school*: National Academies Press.
Brice, A., Shaunessy, E., Hughes, C., McHatton, P., & Ratliff, M. (2008). What language discourse tells us about bilingual adolescents: A study of students in gifted programs and students in general education programs. *Journal for the Education of the Gifted, 32*(1), 7-33.
Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*(7), 513-531.
Bronfenbrenner, U., & Morris, P. (1998). The ecology of developmental processes *The Handbook of Child Psychology: Theoretical Models of Human Development* (5th ed., pp. 993-1028). New York: John Wiley & Sons, Inc.
Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? *Qualitative Research, 6*(1), 97-113.
Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42.
Capps, R., Fix, M., Ost, J., Reardon-Anderson, J., & Passell, J. (2004). *The health and well-being of young children of immigrants*: The Urban Institute.
Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics, 9*(03), 247-266.
Carlsen, W. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research, 61*(2), 157-178.
Cazden, C. (1988). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann Educational Books, Inc.
Cazden, C. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.
Cazden, C., John, V., & Hymes, D. (Eds.). (1972). *Functions of language in the classroom*. New York: Teachers College Press.
Chall, J. (1983). *Stages of reading development*. New York: McGraw-Hill
Chang-Wells, G., & Wells, G. (1993). Dynamics of discourse: Literacy and the construction of knowledge. *Contexts for learning: Sociocultural dynamics in children's development*, 58-90.
Chin, C., & Brown, D. (2000). Learning in Science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*(2), 109-138.
Chin, C., & Osborne, J. (2010). Supporting argumentation through students' questions: Case studies in science classrooms. *Journal of the Learning Sciences, 19*(2), 230-284.
Chinn, C., Anderson, R., & Waggoner, M. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly, 36*(4), 378-411.
Christie, F. (2002). *Classroom discourse analysis: A functional perspective*. New York: Continuum.
Cirino, P., Pollard-Durodola, S., Foorman, B. R., Carlson, C., & Francis, D. (2007). Teacher characteristics, classroom instruction, and student literacy and language outcomes in bilingual kindergartners. *The Elementary School Journal, 107*(4), 341-364.
Coffey, A., & Atkinson, P. (1996). *Making sense of qualitative data: Complementary research strategies*. Thousand Oaks, CA: Sage Publications.

Connor, C., Morrison, F., Fishman, B., Giuliani, S., Luck, M., Underwood, P., et al. (2011). Testing the impact of child characteristics × instruction interactions on third graders' reading comprehension by differentiating literacy instruction. *Reading Research Quarterly, 46*(3), 189-221.

Connor, C., Morrison, F., & Petrella, J. (2004). Effective reading comprehension instruction: Examining child x instruction interactions. *Journal of Educational Psychology, 96*(4), 682-698.

Connor, C., Morrison, F., & Slominski, L. (2006). Preschool instruciton and children's emergent literacy growth. *Journal of Educational Psychology, 98*(4), 665-689.

Connor, C., Morrison, F. J., & Underwood, P. S. (2007). A second chance in second grade: The independent and cumulative impact of first-and second-grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading, 11*(3), 199 - 233.

Creswell, J., & Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2 ed.). Thousand Oaks, CA: Sage Publications, Inc.

Crookes, G. (1990). The utterance, and other basic units for second language discourse analysis. *Applied Linguistics, 11*(2), 183-199.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222-251.

Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press.

Cutting, L., & Scarborough, H. (2006). Predition of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehenison is measured. *Scientific Studies of Reading, 10*(3), 277-299.

Darling-Hammond, L., & Bransford, J. (2005). *Preparing teachers for a changing world: What teachers should learn and be able to do*: Jossey Bass.

Dickinson, D., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E., & Poe, M. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95*(3), 465-481.

Dickinson, D., & Porche, M. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development 82*(3), 870-886.

Dickinson, D., & Tabors, P. (Eds.). (2001). *Beginning literacy with language: Young children learning at home and school*. Baltimore: Paul H. Brookes Publishing Co.

Dillon, J. (1994). *Using discussion in classrooms*. Buckingham, England: Open University Press.

Duffy, B., & Roehler, L. (1982). Direct instriction of comprehension: What does it really mean? *Reading Horizons, 23*, 35-40.

Duke, N. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35*(2), 202-224.

Duke, N., & Carlisle, J. (2011). The development of comprehension. In M. Kamil, D. Pearson, E. Moje & P. Afflerback (Eds.), *Handbook of Reading Research* (Vol. 4, pp. 199-228). New York: Routledge.

Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in the Teaching of English, 23*(1), 4-29.

Elizabeth, T., Anderson, T., Snow, E., & Selman, R. (2012). Academic discussions: An analysis of instructional discourse and an argument for an integrative assessment framework. *American Educational Research Journal*.

Enders, C., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods, 12*(2), 121-138.

Erickson, F., & Mohatt, G. (1982). Cultural organization of participation structures in two classrooms of Indian students. In G. Spindler (Ed.), *Doing the ethnography of schooling: Educational anthropology in action* (pp. 132–174). Austin, TX: Holt, Rinehart and Winston.

Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*: Routledge.

Fitzgerald, J. (1995). English-as-a-second-language learners' cognitive reading processes: A review of research in the United States. *Review of Educational Research, 65*(2), 145-190.

Flanders, N. (1970). *Analyzing teaching behavior*. Reading, MA: Addison-Wesley.

Forman, E., & Cazden, C. (1994). Exploring Vygotskian perspectives in education: The cognitive value of peer interaction. *Theoretical models and processes of reading, 4*, 155-178.

Fraenkel, J., & Wallen, N. (2006). *How to design and evaluate research in education* (6 ed.). New York: McGraw Hill.

Francis, D., Snow, C., August, D., Carlson, D., Miller, J., & Iglesias, A. (2006). Measures of reading comprehension: A latent variable analysis of the Diagnostic Assessment of Reading Comprehension. *Scientific Studies of Reading, 10*(3), 301-322.

Franke, M., Webb, N., Chan, A., Ing, M., Freund, D., & Battey, D. (2009). Teacher Questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education, 60*(4), 380-392.

Gamoran, A., & Nystrand, M. (1991). Background and instructional effects on achievement in eighth-grade English and social studies. *Journal of Research on Adolescence, 1*(3), 277-300.

Gardiner, J. (1980). *Stone Fox*. New York: Harper Collins.

Gee, J. (2001). Reading as situated language: A sociocognitive perspective. *Journal of Adolescent & Adult Literacy, 44*(8), 714-725.

Gee, J. (2005). *An introduction to discourse analysis: Theory and method* (2 ed.). New York: Routledge.

Gee, J., & Green, J. (1998). Discourse analysis, learning, and social practice: A methodological study. *Review of research in education, 23*(1), 119-169.

Gersten, R., & Carnine, D. (1986). Direct instruction in reading comprehension. *Educational Leadership*, 70-78.

Geva, E. (2006). Second-language oral proficiency and second-language literacy. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language minority children and youth* (pp. 123-139). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine de Gruyter.

Goldenberg, C. (2011). Reading instruction for English language learners. In M. Kamil, D. Pearson, E. Moje & P. Afflerback (Eds.), *Handbook of Reading Research* (Vol. 4, pp. 684-710). New York: Routledge.

Goldenberg, C., Gallimore, R., & Reese, L. (2005). Using mixed methods to explore latino children's literacy development. In T. Weisner (Ed.), *Discovering successful pathways in children's development: Mixed methods in the study of childhood and family life* (pp. 21-46). Chicago: The University of Chicago Press.

Graesser, A., & Olde, B. (2003). How does one know whether a person understands a device? The quality of the questions the person asks when the device breaks down. *Journal of Educational Psychology, 95*(3), 524-536.

Greeno, J. (2006). Learning in activity. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 79-96). Cambridge: Cambridge University Press.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and Non-Hispanic students. *Journal of Special Education, 39*(2), 66-85.

Gutiérrez, K., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational Researcher, 32*(5), 19-25.

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education, 5*(2), 93-116.

Hand, B., Wallace, C., & Yang, E. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: Quantitative and qualitative aspects. *International Journal of Science Education, 26*(2), 131-149.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experiences of American children*. Baltimore: Brookes Publishing.

Hart, B., & Risley, T. (2003). The early catastrophe. *American Educator*, 1-6.

Heath, S. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11*(1), 49-76.

Heath, S. (1983). *Ways with words: Language, life, and work in communities and classrooms*: Cambridge University Press.

Herrenkohl, L., & Mertl, V. (2010). *How students come to be, know, and do: A case for a broad view of learning*. New York, NY: Cambridge University Press.

Hill, H., Charalambous, C., & Kraft, M. (2012). When rater reliability is not enough: Teacher observations systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56-64.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development, 62*, 782-792.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55-88.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127-160.

Hufferd-Ackles, K., Fuson, K., & Sherin, M. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*(2), 81-116.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236-248.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*(3), 337-374.

Indrisano, R., & Chall, J. (1999). Literacy development. In R. L. Mosher, D. J. Youngman & J. M. Day (Eds.), *Human development across the life span*. Westport, CT: Praeger.

Johnson, R., & Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14-25.
Joshi, R. (2005). Vocabulary: A critical component of comprehension. *Reading & Writing Quarterly, 21*(3), 209-219.
Juel, C., & Minden-Cupp, C. (2000). Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly, 35*(4), 458-492.
KewalRamani, A., Gilbertson, L., Fox, M. A., & Provasnik, S. (2007). *Status and trends in the education of racial and ethnic minorities (NCES 2007-039)*. Washington, DC: National Center for Educational Statistics, Institute of Education Sciences, US Department of Education.
Kintsch, W. (2004). The construction-integration model of text comprehension and its implications for instruction. *Theoretical models and processes of reading*, 1270-1328.
Klingner, J., & Vaughn, S. (2000). The helping behaviors of fifth graders while using collaborative strategic reading during ESL content classes. *Tesol Quarterly, 34*(1), 69-98.
Krashen, S., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon.
Landis, J., & Koch, G. (1977a). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.
Lawrence, J., & Snow, C. (2011). Oral discourse and reading. In M. Kamil, D. Pearson, E. Moje & P. Afflerback (Eds.), *Handbook of Reading Research* (Vol. 4, pp. 320-337). New York: Routledge.
Lee, J., Grigg, W., & Donahue, P. (2007). *The nation's report card: Reading 2007 (NCES 2007-496)*. Washington, DC: National Center for Educational Statistics, Institute of Educational Sciences, US Department of Education.
Lerman, S. (2001). Cultural, discursive psychology: A sociocultural approach to studying the teaching and learning of mathematics. *Educational Studies in Mathematics, 46*, 87-113.
Lesaux, N., & Geva, E. (2006). Synthesis: Development of literacy in language-minority students. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: A report of the national literacy panel on language-minority children and youth* (pp. 53-74). Mahwah, NJ: Lawrence Erlbaum Associates.
Lin, T., Anderson, R., Hummel, J., Jadallah, M., Miller, B., Nguyen-Jahiel, K., et al. (2012). Children's use of analogy during collaborative reasoning. *Child Development, 83*(4), 1429-1443.
Linnell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives*. Amsterdam: John Benjamins.
Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences, 379*(1), 259-278.
Lotman, Y. (1988). Text within a text. *Journal of Russian and East European Psychology, 26*(3), 32-51.
Luke, D. (2004). *Multilevel modeling*. Thousand Oaks: Sage Publications.
Mackey, A., & Philip, J. (1998). Conversational interactoin and second language development: Recasts, responses, and red herrings? *The Modern Language Journal, 82*(3), 338-356.
Maloch, B. (2002). Scaffolding student talk: One teacher's role in literature discussion groups. *Reading Research Quarterly, 37*(1), 94-112.

Martin, J. (1992). *English text: System and structure*. Philadelphia, PA: John Benjamins, B.V.
McCrudden, M., Magliano, J., & Schraw, G. (2010). Exploring how relevance instructions affect personal reading intentions, reading goals and text processing: A mixed methods study. *Contemporary Educational Psychology, 35*(4), 229-241.
McElhone, D. (2012). Tell us more: Reading, comprehension, engagement, and conceptual press discourse. *Reading Psychology, 33*(6), 525-561.
McGuinness, D. (2005). *Language development and learning to read: The scientific study of how language development affects reading skill.* Cambridge, MA: The MIT Press.
McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203-229.
McQuillan, P. (1998). *Educational opportunity in an urban American high school: A cultural analysis*. Albany: State University of New York Press.
Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press
Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British educational research journal, 25*(1), 95-111.
Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass Inc. Publishers.
Michaels, S., O'Connor, M. C., & Resnick, L. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education, 27*, 283-297.
Miles, M., & Huberman, A. (1994). *Qualitative data analysis: An expanded sourcebook* (2 ed.). Thousand Oaks, CA: Sage Publications, Inc.
Moje, E. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent & Adult Literacy, 52*(2), 96-107.
Moje, E., Collazo, T., Carrillo, R., & Marx, R. (2001). "Maestro, what is 'quality'?": Language, literacy, and discourse in project-based science. *Journal of Research in Science Teaching, 38*(4), 469-498.
Montecillo Leider, C., Proctor, C., Silverman, R., & Harring, J. R. (2013). Exploring the role of vocabulary depth, cross-linguistic transfer, and types of reading measures on the reading comprehension of Latino bilinguals in elementary school. *Reading & Writing Quarterly*.
Morse, J. (2010). Procedures and practice of mixed method design: Maintaining control, rigor, and complexity. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social & behavioral research* (2 ed.). Thousand Oaks, CA: Sage Publications, Inc.
Murphy, P., Wilkinson, I., Soter, A., Hennessey, M., & Alexander, J. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology, 101*(3), 740-764.
Nagy, W., & Scott, J. (2004). Vocabulary processes. In R. Ruddell & N. Unrau (Eds.), *Theoretical models and processes of reading* (5 ed., pp. 269-284). Newark, DE: Internation Reading Association.
Nassaji, H. (2002). Schema theory and knowledge-based processes in second language reading comprehension: A need for alternative perspectives. *Language Learning, 52*(2), 439-481.
Nassaji, H., & Wells, G. (2000). What's the use of 'triadic dialogue'?: An investigation of teacher-student interaction. *Applied Linguistics, 21*(3), 376-406.

Nathan, M., & Knuth, E. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21*(2), 175-207.
Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356.
National Council of Teachers of English. (2006b). *NCTE principles of adolescent literacy reform: A policy brief*. Urbana, IL: National Council of Teachers of English.
National Governors Association Center for Best Practices Council of Chief State School Officers. (2010). *Common Core State Standards*.
NCELA. (2008). National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs. Retrieved February 22, 2009, from http://www.ncela.gwu.edu/expert/faq/08leps.html
NCES. (2013). U.S. Department of Education, National Center for Education Statistics. (2013). The Condition of Education 2013 (NCES 2013-037), English Language Learners., from http://nces.ed.gov/fastfacts/display.asp?id=96
Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning, 51*(4), 719-758.
NIH. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for instruction.*: National Institute of Child Health and Human Development.
Nystrand, M. (2006). Research on the role of classroom discourse as it affects reading comprehension. *Research in the Teaching of English, 40*(4), 392-412.
Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English, 25*(3), 261-290.
Nystrand, M., Gamoran, A., Kachur, R., & Prendergast, C. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Teachers College Press
Nystrand, M., Wu, L., Gamoran, A., Zeiser, S., & Long, D. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35*(2), 135-198.
O'Connor, M. (2001). "Can any fraction be turned into a decimal?" A case study of a mathematical group discussion. *Educational Studies in Mathematics, 46*(1), 143-185.
O'Connor, M., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology and Education Quarterly, 24*(4), 318-335.
Orellana, M., & Reynolds, J. (2008). Cultural modeling: Leveraging bilingual skills for school paraphrasing tasks. *Reading Research Quarterly, 43*(1), 48-65.
Ouellette, G. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554 - 566.
Palincsar, A. (1998). Social constructivist perspectives on teaching and learning. *Annual review of psychology, 49*, 345-375.
Palinscar, A., & Brown, A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and instruction, 1*(2), 117-175.
Pea, R. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47-87). Cambridge: Cambridge University Press.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. Kamil, P. Mosenthal, P. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 545-562). Mahwah, NJ: Erlbaum.

Pressley, M., El-Dinary, P., Gaskins, I., Schuder, T., Bergman, J., Almasi, J., et al. (1992). Beyond direct explanation: Transactional instruction of reading comprehension strategies. *The Elementary School Journal, 92*(5), 513-555.

Pressley, M., Wharton-McDonald, R., Allington, R., Block, C., Morrow, L., Tracey, D., et al. (2001). A study of effective first-grade literacy instruction. *Scientific Studies of Reading, 5*(1), 35-58.

Proctor, C., Dalton, B., Uccelli, P., Biancarosa, G., Mo, E., Snow, C., et al. (2011). Improving comprehension online: Effects of deep vocabulary instruction with bilingual and monolingual fifth graders. *Reading & Writing Quarterly, 24*, 517-544.

Proctor, C., Silverman, R., Harring, J., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish-English bilingual children in elementary school. *Reading and Writing, 25*(7), 1635-1664.

RAND. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Corporation.

Raudenbush, S., & Bryk, A. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59*(1), 1-17.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2 ed.). Thousand Oaks, CA: Sage Publications, Inc.

Renshaw, P. (2004). Dialogic learning, teaching, and instruction: Theoretical roots and analytical frameworks. In J. Van Der Linden & P. Renshaw (Eds.), *Dialogic learning: Shifting perspectives to learning, instruction, and teaching*. Boston, MA: Kluwer Academic Publishers.

Reznitskaya, A., Anderson, R., McNurlen, B., Nguyen, K., Archodidou, A., & Kim, S. (2001). Influence of oral discussion on written argumentation. *Discourse Processes, 32*(2-3), 155-175.

Riley, J. (1986). The effects of teachers' wait-time and knowledge comprehension questioning on science achievement. *Journal of Research in Science Teaching, 23*(4), 335-342.

Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In J. Wertsch, P. del Rio & A. Alvarez (Eds.), *Sociocultural stuides of the mind* (pp. 139-163). Cambridge, UK: Cambridge University Press.

Rosenblatt, L. (1994). The transactional theory of reading and writing. In R. Ruddell, M. Ruddell & H. Singer (Eds.), *Theoretical models and processes of reading* (4 ed., pp. 1057-1092). Newark, DE: International Reading Association.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*(4), 479-530.

Ruddell, R., & Unrau, N. (2004). Reading as a meaning-construction process: The reader, the text and the teacher. In R. Ruddell & N. Unrau (Eds.), *Theoretical Models and Processes of Reading* (5 ed.). Newark, DE: International Reading Association.

Ryan, P. (2000). *Esperanza rising*. New York: Scholastic Inc.

Rydland, V., Grover, V., & Lawrence, J. (2013). The second-language vocabulary trajectories of Turkish immigrant children in Norway from ages five to ten: The role of preschool talk

exposure, maternal education, and co-ethnic concentration in the neighborhood. *Journal of Child Language, 41*, 352-381.
Saunders, W., & Goldenberg, C. (1999). Effects of instructional conversations and literature logs on limited-and fluent-English-proficient students' story comprehension and thematic understanding. *The Elementary School Journal, 99*(4), 277-301.
Sawyer, R. (2006). Analyzing collaborative discourse. In R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 187-204). New York: Cambridge University Press.
Scardamalia, M., & Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *Journal of the Learning Sciences, 1*(1), 37-68.
Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. *Cambridge handbook of the learning sciences*, 97-118.
Schleppegrell, M. J. (2006). The challenges of academic language in school subjects. In I. Lindberg & K. Sandwall (Eds.), *Spraket och kunskapen: att lara pa sitt andrasprak i skola och hogskola* (pp. 47-69). Sweden: Goteborgs Universitet.
Scientific Software Development. (2011). ATLAS.ti 6.2. Berlin, Germany.
Scott, J., Jamieson-Noel, D., & Asselin, M. (2003). Vocabulary instruction throughout the day in twenty-three Canadian upper-elementary classrooms. *The Elementary School Journal, 103*(3), 269-286.
Semel, E., Wiig, E., & Secord, W. (2003). *Clinical evaluation of language fundamentals - 4*. San Antonio, TX: PsychCorp/Harcourt.
Sharpe, T. (2008). How can teacher talk support learning? *Linguistics and Education, 19*(2), 132-148.
Shavelson, R. (1996). *Statistical reasoning for the behavioral sciences* (3 ed.). Needham Heights, MA: Allyn & Bacon.
Silverman, R., & Crandell, J. (2010). Vocabulary practices in prekindergarten and kindergarten classrooms. *Reading Research Quarterly, 45*(3), 319-340.
Silverman, R., Proctor, C., Harring, J., Doyle, B., Mitchell, M., & Meyer, A. (2013). Teachers' instruction and students' vocabulary and comprehension: An exploratory study with English monolingual and Spanish–English bilingual students in grades 3–5. *Reading Research Quarterly, 49*(1), 31-60.
Sim, J., & Wright, C. C. (2005). The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*(3), 257-268.
Sinclair, J., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press, Inc.
Slavin, R., Cheung, A., Groff, C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best-evidence synthesis *Reading Research Quarterly, 43*(3), 290-322.
Slavin, R., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary grades: A best-evidence synthesis. *Review of Educational Research, 79*(4), 1391-1466.
Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the Classroom Observations of Student–Teacher Interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly, 27*(2), 316-328.

Snow, C. (1972). Mothers' speech to children learning language. *Child development*, 549-565.
Snow, C. (1991). The theoretical basis for relationships between language and literacy in development. *Journal of Research in Childhood Education, 6*(1), 5-10.
Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*: RAND.
Snow, C., Barnes, W., Chandler, J., Goodman, I., & Hemphill, L. (1991). *Unfulfilled expectations: Home and school influences on literacy*. Cambridge, MA: Harvard University Press.
Snow, C., Burns, M., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washingon, DC: National Academy Press.
Snow, C., Griffin, P., & Burns, M. (2005). *Knowledge to support the teaching of reading: Preparing teachers for a changing world*: Jossey-Bass.
Snow, C., Porche, M., Tabors, P., & Harris, S. (2007). *Is literacy enough? Pathways to academic success for adolescents*. Baltimore, MD: Brookes Publishing Company.
Soter, A., Wilkinson, I., Murphy, P., Rudge, L., Reninger, K., & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research, 47*(6), 372-391.
Speare, E. (1983). *The sign of the beaver*. New York: Bantam Doubleday Dell Books for Young Readers.
Stahl, S., & Fairbanks, M. (1986). The effects of vocaulary instruction: A model-based meta-analysis. *Review of Educational Research, 56*, 72-110.
Stake, R. (2006). *Multiple case study analysis*. New York, NY: The Guilford Press.
StataCorp. (2011). *Stata Statistical Software: Release 12*. College Station, TX: StataCorp LP.
Stevens, R., Slavin, R., & Farnish, A. (1991). The effects of coooperative learning and direct instruction in reading comprehension strategies on main idea identification. *Journal of Educational Psychology, 83*(1), 8-16.
Strizek, G. A., Pittsonberger, J. L., Riordan, K. E., Lyter, D. M., & Orlofsky, G. F. (2006). *Characteristics of Schools, Districts, Teachers, Principals, and School Libraries in the United States: 2003-04 Schools and Staffing Survey (NCES 2006-313)*. Washington, DC: US Department of Education.
Swain, M. (2005). The output hypothesis: Theory and research. *Handbook of research in second language teaching and learning*, 471–483.
Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371-391.
Tashakkori, A., & Creswell, J. (2007). Editorial: Exploring the nature of research questions in mixed methods research. *Journal of Mixed Methods Research, 1*(3), 207-210.
Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research* (pp. 3-50). Thousand Oaks: Sage Publications.
Tharp, R. G. (1982). The effective instruction of comprehension: Results and description of the Kamehameha Early Education Program. *Reading Research Quarterly, 17*(4), 503-527.
Turner, J., & Meyer, D. (2000). Studying and understanding the instructional contexts of classrooms: Using our past to forge our future. *Educational Psychologist, 35*(2), 69-85.

Turner, J., Meyer, D., Midgley, C., & Patrick, H. (2003). Teacher discourse and sixth graders' reported affect and achievement behaviors in two high-mastery/high-performance mathematics classrooms. *The Elementary School Journal, 103*(4), 357-382.
UCLA. Analyzing Correlated (Clustered) Data. *UCLA: Statistical Consulting Group.* from http://statistics.ats.ucla.edu/stat/stata/library/cpsu.htm. Retrieved August 10, 2013
Vygotsky , L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: The MIT Press.
Waggoner, M., Chinn, C., Yi, H., & Anderson, R. (1995). Collaborative Reasoning about stories. *Language Arts, 72*(8), 582-589.
Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65*(2), 606-621.
Watts, S. M. (1995). Vocabulary instruction during reading lessons in six classrooms. *Journal of Reading Behavior, 27*(3), 399-424.
Weizman, Z., & Snow, C. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental psychology, 37*(2), 265-279.
Wells, G. (1990). Talk about text: Where literacy is learned and taught. *Curriculum Inquiry, 20*(4), 369-405.
Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. New York: Cambridge University Press.
Wells, G. (2000). Dialogic inquiry in Education: Building on the legacy of Vygotsky. In C. Lee & P. Smagorinsky (Eds.), *Vygotskian perspectives on literacy research* (pp. 51-85). New York, NY: Cambrigde University Press.
Wells, G., & Arauz, R. (2006). Dialogue in the classroom. *Journal of the Learning Sciences, 15*(3), 379-428.
Wertsch, J. (1998). *Mind as action*. New York, NY: Oxford University Press.
Wiebe Berry, R., & Kim, N. (2008). Exploring teacher talk during mathematics instruction in an inclusion classroom. *The Journal of Educational Research, 101*(6), 363-378.
Wilkinson, I., & Son, E. (2011). A dialogic turn in research on learning and teaching to comprehend. In M. Kamil, D. Pearson, E. Moje & P. Afflerback (Eds.), *Handbook of Reading Research* (Vol. 4, pp. 359-387). New York: Routledge.
Wolf, M., Crosson, A., & Resnick, L. (2004). Classroom talk for rigorous reading comprehension instruction. *Reading Psychology, 26*(1), 27-53.
Woodcock, R., Munoz-Sandoval, A., Reuf, M., & Alvarado, C. (2005). *Woodcock-Muñoz Language Survey-Revised*. Itasca, IL: Riverside Publishing Company.
Yin, R. (2009). *Case study research: Design and methods* (4 ed.). Thousand Oaks, CA: Sage Inc.
Zhang, J., Anderson, R., & Nguyen-Jahiel, K. (2013). Language-rich discussions for English language learners. *International Journal of Educational Research, 58*, 44-60.

Appendix A: Coding Scheme

Definitions:

Instructional Episode: A coherent learning activity centering around a particular objective or purpose; "bounded sequences of actions which are somehow coherent internally" (Linell, 1998, p.xiv). A new episode is marked when the T addresses a new objective (new topic). Each episode is marked with Global Features (GF) codes after reading through the lesson. Some lessons have one episode; most lessons range between having 2 to 4 instructional episodes in an hour-long lesson.

Utterance: The smallest segment of intact talk by a student or teacher that conveys an idea. Eg. "Nice work." "Why do you think that?" Teacher and student utterances coded with Utterance Features (UF) codes, including Question Quality (QU) codes (if utterance is a query). Some utterances have 2-3 different codes applied to them, including Teacher questions codes that must have an accompanying QU code.

Segment: A group of coherent utterances that focus on one field/topic. Eg. All students' responses related to a teacher's question, or a student's response to a T's uptake question before the teacher turns to another student with a different question.

Global Features (GF) of lessons: After each instructional episode is identified in each transcript, each episode is coded once for its global features: who manages the activity in the episode (GF_TM or GF_SM), what instructional materials are used in each episode (GF_Mat), what kind of instructional activity (e.g. pattern of talk) the class engaged in (GF_Gen), and what ELA content topic, or field, is primarily addressed in the episode (GF_Fie).

Utterance Level Features (UF) of lessons: Each teacher or student utterance that is related to instruction (and not procedural, managerial or rhetorical) is coded for 12 codes (not including st questions) relating to dialogic instruction. These are types of teacher questions, question quality, types of teacher evaluations, L1 use, and teacher or student extended explanations.

Coding Notes:

*Don't code for student or teacher procedural, rhetorical, or managerial comments/questions. Coding will be episodes and utterances that forward instruction of some content, not manage how the activity is enacted. For example, if a teacher asks "Amy, what did you just say?" this shouldn't be marked as an Uptake; these clarification questions don't further the conversation or the lesson necessarily, and seem more procedural. However, when the teacher asks if the class heard a speaker, or attends to the nature of speaking to the group, this is to be coded because it explicitly attends to the importance of dialogue between actors (Uptake if posed as a question, or as Contiguity if a statement).

*Don't code repeated questions if the repeated question had no scaffolding beyond the first question.

*Don't code for teacher or student read alouds, for student repetitions of teacher output verbatim, or for pronunciation help from the teacher when a student is reading aloud. These instructional acts forward the lesson, but these are not dialogic segments of instruction.

Code	Code Name	Definition	References	Exemplars
Global features (GF) of lessons				
GF_TM GF_SM	Management of discourse	Who manages the instructional progress/flow of the lesson: Teacher(s) or students?	Chinn, Anderson &	

	[to be coded after transcript is read and episodes are identified]	Teacher or students drive the progress of the instructional episode by controlling turn-taking to advance the lesson for student learning; Teacher or students are responsible for "focusing attention on learning" (Connor et al., 2007, p.209)	Waggoner, 2001 Connor et al., 2007	
GF_Mat_[name] GF_Mat_Basal GF_Mat_Unabr GF_Mat_GO GF_Mat_Info GF_Mat_Other GF_Mat_StTxt GF_Mat_Tstg GF_Mat_Video GF_Mat_WkhFF GF_Mat_WkhVoc	Instructional Materials *[to be coded after transcript is read and episodes are identified]*	Materials used in the literacy/ELA lesson: *Basal: basal readers, literature anthology, or leveled readers (Open Court, Houghton-Mifflin Reading, Rigby Literacy Student Readers) *Unabridged texts: unabridged literature of any length, incl. short stories, poems *Graphic organizers (student- or teacher-made) *Informational texts: science materials, magazines, newspapers, abridged and unabridged *Other: special projects *Student Text: student-written texts (journals, practice test responses, short writing activities, etc.) *Testing: spelling, vocabulary tests, explicit test-prep materials, materials in test-like format (m/c, open response) *Video *Worksheets, form-focused (grammar, spelling, punctuation, capitalization) *Worksheets, vocabulary		*Anthology:* T to class: …comes from the book, Facing Your Fear. This story comes from a unit in the book called *Give It All You've Got*… excerpt from *The Fear Place* by Phyllis Reynolds Naylor. *Form-focused worksheets*: T to class: Take out your Tuesday Daily language practice homework.
GF_ Gen_[name] GF_Gen_1x1Rdg GF_Gen_DI GF_Gen_Ind GF_Gen_Q&A	Literacy activity/genre *[to be coded after transcript is read and*	Activities sts engaged in, or, "genre," a "staged purposive activity undertaken to accomplish some goal" *1x1 Reading: individual student working with teacher or aid on a reading activity	Christie, 2002, p. 21	*Direct instruction:* T holds up a card, with 'versatile'. T reads definition of word aloud from worksheet. T gives an example of versatile. T holds up second word, 'comforting'. Students echo the word. T gives a definition of comforting. T calls on student to give

GF_Gen_RA GF_Gen_ SmGrTxtTlk GF_Gen_Class TxtTlk	*episodes are identified]*	*Direct Instruction: instructional talk dominated by teacher explanations, usually to present new content *Individual seat work: students working alone on worksheets, reading, writing, testing *Q&A: teacher poses questions and students share answers in whole class or small groups (usually class is going over previous work) *Read Aloud: majority of instructional episode is spent accomplishing a reading; main focus is on reading aloud *Small group Text-based Talk: main attention of episode is on text comprehension, applying reading strategies or background knowledge; talking about the text in groups of 2-6 students *Whole-class Text-based Talk: same as above, with majority of students		definition of comforting. T holds up third word, 'fleet'. T reads definition aloud for fleet. *SmGrp Text Talk:* T: All right, boys and girls, we're going to pass out to you each a question that goes along with author's purpose and audience. Today we're going to work in groups, so we're going to be doing a lot of this on your own.
GF: Fie. [topic] GF_Fie_Bckgrd GF_Fie_FF GF_Fie_Mod GF_Fie_RdgStrat GF_Fie_Rev GF_Fie_SpEvent GF_Fie_Tst GF_Fie_TxtCom p GF_Fie_Vocab GF_Fie_Wtg	Field (topic) *[to be coded after transcript is read]*	Field of discourse; what is going on; the topic/content under discussion/instruction: *Background: building or sharing background knowledge in conjunction with text class will read/has read *Form-focused work: capitalization, punctuation, syntax & grammar, spelling, handwriting *Modeling: T models academically appropriate work for majority of episode *Reading Strategy: T is explicit about a reading strategy(ies) in the context of a text (connecting to prior knowledge, inferring, predicting, functions of text features, etc.) *Review of previous work, usually setting up for next learning activity *Special Event: holiday project, presentation in auditorium	Christie, 2002	*Modeling:* T writes 'summary' on chart and asked students to get out a half sheet of paper, T writes first and last main idea and then tells students to write own summary using map (graphic organizer) *Reading Strategy:* T: We're going to review text features. We've already talked a little bit about them. Text features basically are any other things we see when we're reading that are different from just plain, boring, normal letters and words. *Writing:*

		*Testing: testing (quizzes, DRAs, etc.) or test preparation using testing materials (test taking strategies, practice standardized tests, etc.) *Text Comprehension: main focus of episode is on text comprehension *Vocabulary: main focus of episode is on vocabulary comprehension, sometimes connected to a text *Writing: main focus of episode is on student writing, including free writing, writing about text, collaborative writing, or research and notetaking for writing (but not writing for tests)		T: Go ahead and take out your rough draft that we've been working on. It should be on your clipboard.
Utterance level features	Dialogic Indicators	Description	References	Exemplars
UF_Q_ AQ UF_Q_QQ UF_Q_TQ UF_Q_Up	Authenticity of Teacher Questions *(accompanied by a QU code (below))* **Authentic** or **Quasi-authentic** or **Test/Display** or **Uptake**	**Authentic teacher questions* – Teachers questions do not have a prespecified answer that the teacher is seeking and do not show uptake of previously contributed ideas; Indeterminate number of possible answers; Teacher seeking more student answers in certain contexts (e.g. *What else do you think? Any other ideas?)* -AQ can be accompanied by a QU: Rec code- see examples **Quasi-authentic*: Teacher questions that offer sts "some latitude in answering" but have a finite range of possibilities; some definitions/answers are acceptable; this includes most questions about text: the text necessarily limits the answers a student can provide (so not usually an AQ)	Applebee et al., 2003, p. 700 Nystrand et al., 2003, p.145 Gamoran & Nystrand, 1991, p.289 Boyd & Rubin, 2006 McNeill & Pimentel 2010	*Authentic Questions:* T: Can you think of some ways that people might deal with that? [AQ,Gen] T: What do you think? T: What else did your group write down? [AQ, Rec] T: Dasha, do you want to add something to that? [AQ, Rec] T: When in the real world does that happen? Think and share. T: Did you like the ending, not like the ending? You surprised about anything? [Eval] *Quasi-authentic Questions:* T: What are some of the corrections we should be making? [QQ, Rec or Ana] T: Comfort. What's comfort? T: What happened in this part of the story? [QQ,Rec] T: Okay, so who can give me a detail about animals migrating [from the text]?

		**Test/Display questions*: Teacher is presumed to know answer (or class is presumed to know answer based on earlier talk in class), and question has only one possible correct answer. These are typically low-cognitive quality, but not always (e.g. m/c reading comprehension questions can be QU: Gen or Ana) **Uptake questions*: Indicator of reciprocity of talk on the topic, Uptake is marked by teacher questions that incorporate a previous response; teacher asks students follow-up questions to pursue points or lines of inquiry introduced by students; Uptake is a question that incorporates a previous answer/statement made by a student, not a film, text, a previous T question, or worksheet. -Uptake is often marked by use of pronouns: How did *it* work? And what caused *this*? And *then* what happened? -Uptake questions can incorporate a previous st question if the T restates this question for the whole class, positioning this question as central to instruction for that moment - called this "connected": "dialogic interactions that support, refute, restate, or ask a clarifying question about a previous idea" (McNeill & Pimentel, 2010, p. 212) -called "contingent" questions that extend or respond to preceding utterances (Boyd & Rubin, 2006, p.153) *-Note: students always exhibit uptake in their questions if they attend to the discussion, while teacher uptake varies*		T: Show me or tell me what surprised would look like. T: Do you remember why you wrote that? [QQ, Ana] *Test/Display Questions:* T: What's the other meaning for stationery? A totally different meaning. [TQ,Rec] T: Why do we capitalize the J? Because it is, what? [TQ, Rec] T: Can a plant thrive without water, yes or no?[TQ,Rec/Gen] T: If I were to say literal or figurative language, which one is that? What are they [the book] using? [TQ,Rec or Gen] T: What's the simile in the first sentence? [TQ,Ana] T reads out m/c question aloud [TQ,Gen or Ana] *Uptake Questions*: T: There could be. But what does that have to do with the dark? What might that have to do with the dark? T: Why? Tell me more. T: You want to add anything to the Achilles' story [*that a student just told*]? [Up,Rec] T: Did you hear that everyone? Hold on, can you hear Tyler? [Up,Rec] T: Ok, show me or tell me what you would look like if you were surly. They're acting that way because they are very--[Up,Rec].St: Surly. T: And what's a synonym (for surly)? [QQ,Rec] T: Everybody here says Willy will win? Why? [Up,Ana] St: even though his eye is

		<u>Note:</u> *Code for the main emphasis of the question, rather than a "tag" question that may be involved: eg. "Do you remember why you wrote that?" posed when sts were sharing their writing. "Do you remember?" is not the main thrust of this question; rather, "why (did) you write that?" is what gets followed up on, so the T's intent was the latter in this context.* <u>Note:</u> *Procedural, rhetorical, managerial questions not coded*		hurt, Willy still knows the track really well [StExp]. T: So that would work to his advantage? [Up, Rec]
UF_TEvalH UF_TEvalL	High or low level teacher evaluation of student responses	*High-level evaluations:* when a teacher acknowledges that a student has contributed something new that changes or modifies the topic or direction of discussion through the student's response/question/ contribution to a topic. Teacher uses student's response/contribution to further an explanation, or further the development of the topic or lesson. This can be done by "revoicing" by "recontextualizing" or "recasting" a student's contribution to align it with academic task, including the facilitation of students' alignment with "intellectual socialization" including academic language. *Low-level evaluations*: when teacher indicates no or shallow evaluation of student response. The teacher's response to a student utterance(s) may be praise or express sympathy, but the teacher's response does not forward the lesson topic. Teacher repeats student response verbatim but doesn't expand.	Nystrand & Gamoran, 1991 Nystrand et al., 2003, p.146. Nystrand et al. 1997, p. 32 O'Connor & Michaels, 1993 Cazden, 1998	*TEvalHigh:* St: …That the animals—their streams and lakes are unfreezing. T: Right, they're melting, okay. T: 'Troubled' really doesn't mean when you're in trouble. I may be troubled by a situation that happened at recess where some of my students were fighting, and I feel troubled about that. So what do you think troubled means? [QQ T: It was so hot outside, it was raining popcorn because of what? St: A heat wave T: hmmm [TEvalL] St: The corn got so hot. T: Good. The corn growing in the field got so hot it popped into popcorn…[TEvalH] T: Should it 'mortify' someone to make a sandwich? Sts: No! T: You should not be embarrassed to make a sandwich. *TEvalLow:*

		Note: Code for T utterances only, as sts don't evaluate T's questions or utterances		T: Good. Ok. T: That must have been a scary situation. T: Cautious. That's a good word. T: responds only with a 'no' to an incorrect student answer. T: no response to st answer or contribution T: repeats st's response verbatim
UF_UseL1	L1 use	L1 use (mainly Spanish in this sample), by teacher or students, to support student understandings or participation in instructional discourse	(Slavin and Cheung, 2005)	(in field notes): T goes to another st and speaks to him in Spanish to draw their ideas
UF_TExp UF_StExp	Teacher or student explanations nb. Mercer et al, 1999 has a linguistic framework for analysing discussions that defines st explanations as '100 characters when transcribed' that they consider effective indicators of exploratory talk (see also Elizabeth et al. 2012 AERJ article)	*Teacher Explanations:* Teacher's utterance forwards the lesson by explaining concept/topic without student input, without posing questions; more than 2 ideas or 2 lines of transcript. *Student explanations* are elaborations of their ideas, more than 2 lines of transcript with a coherent idea. This is a marker for increased level of student talk. Student writing that is shared aloud with a teacher, peers, or the class, is also considered a student explanation. Student explanations are a proxy for elaborated or extended amounts of student talk/writing/output. Note: *Explanations are not procedural or managerial in content.*	Aukrust , 2007 Soter et al., 2008	*Teacher Explanations:* T: Think about a lot of the fears here. A rollercoaster, the dark, a gross tarantula or spider, snakes. Think about the word control. When we have control over a situation, when you have control over something, not [NOISE OBSCURES] you have a friend over, so it's your house. It's your stuff. You have the control. But then it's a little different when you go to somebody else's house and they're the one who's in charge, so to speak. Think about control and fear. If we can't control the situation, that's a lot of times when kids and adults get a little scared. *Student Explanations:* St: When I watched this show, two of the contestants on the show, when they were chosen, one hyperventilated. The other froze and dropped the phone. St: I ran outside as light as a leaf.

Quality of Teacher Questions		Definition	References	Exemplars
UF_QU_Low UF_QU_High	Cognitive emphasis of teacher questions	Modified from Nystrand et al, 1997, the questions are coded in terms of what thinking they elicit: *Low Cognitive Level of Questions:* These question asks students to <u>Recite known information, Record an answer, and/ or Report information that is presumed knowledge</u>. For example: What's/what did happen? What is this? What does XYZ mean? Fill in the blank questions. The answers to these questions are found in text or in previous T or St utterance/ explanation. Answers draw from routine application of prior knowledge. *High Cognitive Level of Questions:* *<u>Generalization:</u>* What happens? What do I make of what happens? Displays inductive reasoning, building up ideas, takes sts out of the text somewhat. For example, constructing a summary of a text involves deleting trivial and redundant information and constructing superordinate ideas from detailed ideas. [efferent in stance] *<u>Analysis:</u>* Why? Displays deductive reasoning, breaking ideas, arguments down [efferent stance] *<u>Speculation/Prediction:</u>* What might happen [if…]? Asking students to infer possibilities based on	Nystrand et al., 1997, drawing on Applebee (1981), Moffett (1968), Britton et al. (1975) Nystrand et al., 2003 Rosenblatt, 1994	*Low Level Questions:* T: What does that mean to 'face your fear'? T: Have you ever heard somebody saying you have a big mouth? T: Like the ocean, a pool, or just a glass of water? T: Does anyone know where your Achilles' tendon is? T: You want to add anything to the Achilles' story? T: What does 'discover' mean? T: What's another way of saying that? *High Level Questions:* *Generalization:* T: What is the main idea in paragraph 2? T: Give me a one-sentence summary of what we just read. T: Can you think of some ways that people might deal with that? T: And '-ed' at the end of the word tells us about what? T: Can you give me an example of what an animal has to do to survive? *Analysis:* T: What can you infer what happened? Anthony. T: We're going to be asking ourselves how can this word be connected to what Chris Van Allsburg did with the story? T: How do you know? T: Since he felt that he was alone, how did that affect his school? T [reading]: "They stopped a moment ot

		understanding of text. Speculation requires inferring from texts. *<u>Evaluation:</u>* What do you think about XYZ? More personal in response, usually (but not always) at the end of the lesson. When students are asked to make a connection from their background knowledge to a text, the evaluate their own experience for its applicability to the current learning task. [Both Spec and Eval could be efferent or aesthetic in stance] *Procedural, rhetorical, or managerial questions not coded*		listen to the wide blades of grass whisper and squeak in the wind." What form of descriptive language is that? *Speculation/Prediction:* T: I know that one [magazine article] looks interesting, "No More Happy Meals"? I wonder what that's about? T: So I need everyone to turn to Ch.4. It says "Help from a friend." I see that title, and I'm wondering to myself—who this friend is? Who do we think the friend might be? *Evaluation:* T: Okay. But do you think that is the main idea of the whole passage? Is that what the author wants you to take away from this? T: So how do you think the kids feel about their mom?

Appendix B: Code Frequencies of Global Features of Lessons

Table B.1 *Frequencies of Instructional Materials Used across Lessons*

	N	Frequency	Percent	Cumulative percent
Basals	0	62	70.45	70.45
	1	18	20.45	90.91
	2	5	5.68	96.59
	3	2	2.27	98.86
	5	1	1.14	100.00
Unabridged Narrative Texts	0	45	51.14	51.14
	1	24	27.27	78.41
	2	13	14.77	93.18
	3	6	6.82	100.00
Graphic Organizers	0	68	77.27	77.27
	1	13	14.77	92.05
	2	7	7.95	100.00
Informational Texts	0	68	77.27	77.27
	1	14	15.91	93.18
	2	6	6.82	100.00
Other	0	82	93.18	93.18
	1	5	5.68	98.86
	2	1	1.14	100.00
Student Texts	0	62	70.45	70.45
	1	19	21.59	92.05
	2	4	4.55	96.59
	3	3	3.41	100.00
Tests	0	60	68.18	68.18
	1	16	18.18	86.36
	2	6	6.82	93.18
	3	5	5.68	98.86
	4	1	1.14	100.00
Videos	0	84	95.45	95.45
	1	3	3.41	98.86
	2	1	1.14	100.00
Form-focused Worksheets	0	71	80.68	80.68
	1	12	13.64	94.32
	2	4	4.55	98.86
	3	1	1.14	100.00
Vocabulary Worksheets	0	73	82.95	82.95
	1	13	14.77	97.73

	2	2	2.27	100.00

Note: Transcript n = 88; Basals = textbooks, anthologies, leveled readers; Unabridged Narrative Texts = chapter books, short stories, poems; Informational Texts = abridged or unabridged non-narrative/fictional texts; Other = special holiday projects; Student Texts = student written/generated texts; Testing Materials = ELA curriculum-based tests, standardized test-prep materials; Videos = ELA-related video content; FF Worksheets = form-focused worksheet; Vocabulary Worksheets = vocabulary lists usually associated with basals

Table B.2 *Frequencies of Talk Genres used across Transcripts*

	N	Frequency	Percent	Cumulative
1x1 Reading	0	84	95.45	95.45
	1	4	4.55	100.00
Direct	0	75	85.23	85.23
Instruction	1	13	14.77	100.00
	0	41	46.58	46.59
Individual	1	42	47.73	94.32
Work	2	5	5.68	100.00
	0	22	25.00	25.00
	1	39	44.32	69.32
Q&A	2	20	22.73	92.05
	3	5	5.68	97.73
	4	1	1.14	98.87
	5	1	1.14	100.01
	0	38	43.18	43.18
Read Aloud	1	37	42.05	85.23
	2	10	11.36	96.59
	3	3	3.41	100.00
	0	57	64.77	64.77
Small Group	1	21	23.86	88.63
Text Talk	2	6	6.82	95.45
	3	3	3.41	98.86
	4	1	1.14	100.00
	0	73	82.95	82.95
Class	1	12	13.64	96.60
Text Talk	2	3	3.41	100.00

Note: Total number of transcripts=88; 1X1Reading = 1 student reading with 1 teacher or aid; Direct Instruction = majority of talk dominated by teacher explanations; Individual Work = students working individually; Q&A = whole class or small group question & answer pattern with teacher posing questions (e.g. review, vocabulary) and receiving brief answers; Read Aloud = whole class or small group reading of text with teacher; Small Group Text Talk = group of students talking about a text, with or without a teacher; Class Text Talk = whole class talking about text in some depth, attention to text comprehension, applying reading strategies or background knowledge.

Table B.3 *Frequencies of Fields (Topics) of Instruction Addressed across Transcripts*

	N	Frequency	Percent	Cumulative
Background Information	0	81	92.05	92.05
	1	7	7.95	100.00
Form-focused topic	0	68	77.27	77.27
	1	16	18.18	95.45
	2	2	2.27	97.73
	3	1	1.14	98.86
	4	1	1.14	100.00
Modeling	0	83	94.32	94.32
	1	3	3.41	97.73
	2	2	2.27	100.00
Reading Strategies	0	60	68.18	68.18
	1	20	22.73	90.91
	2	7	7.95	98.86
	3	1	1.14	100.00
Review	0	71	80.68	80.68
	1	15	17.05	97.73
	2	2	2.27	100.00
Special Event	0	86	97.73	97.73
	1	1	1.14	98.86
	2	1	1.14	100.00
Test	0	69	78.41	78.41
	1	15	17.05	95.45
	2	4	4.55	100.00
Text Comprehension	0	26	29.55	29.55
	1	38	43.18	72.73
	2	13	14.77	87.5
	3	8	9.09	96.59
	4	2	2.27	98.86
	5	1	1.14	100.00
Vocabulary	0	48	54.55	54.55
	1	27	30.68	85.23
	2	10	11.36	96.59
	3	3	3.41	100.00
Writing	0	52	59.09	59.09
	1	24	27.27	86.36
	2	7	7.95	94.32
	3	5	5.68	100.00

Note: total number of transcripts = 88

Appendix C: Example of a Text-based Discussion

Turn	Speaker	Utterance
1	TEACHER 73:	So, guys. So now, we have Stone Fox, this racer who's raced so many races and won every single one. And then we've got Willy who's never done a race at all, and has swollen eye. Who do you think is going to win?
2	STUDENT:	I think it will be Willy, because Willy's the good guy. [OVERLAPPING VOICES]
3	TEACHER:	OK. Everyone here says Willy? [OVERLAPPING VOICES]
4	TEACHER:	Why? Diana.
5	STUDENT:	Because it's [UNINTELLIGIBLE PHRASE] even though his eye is hurt, sometimes Willy knows the track very well.
6	TEACHER:	Ooh. So that would work to his advantage, wouldn't it? [OVERLAPPING VOICES]
7	TEACHER:	Good job Tiana. Jen, what do you think?
8	STUDENT:	You also need a [UNINTELLIGIBLE PHRASE].
9	STUDENT:	He can still look with his left eye.
10	TEACHER:	He still has his left eye.
11	STUDENT:	Yeah, but he wouldn't see really good.
12	TEACHER:	I know. Why do you think Willy's going to win this?
13	STUDENT:	Because Will's younger.
14	TEACHER:	Yeah? What do you think? Tell me more, Beth.
15	STUDENT:	Now is this who you think will win? [OVERLAPPING VOICES]

The teacher sets up this reciprocal exchange by posing an authentic question "who do you think is going to win?" (turn 1) and pushes students to explain their prediction by asking "why?" (turn 4). This generates a lot of overlapping talk by the students, with some reciprocal exchanges between them, such as when one student responds to another's idea with "Yeah, but he wouldn't see really good" (turn 11) and another student carries the discussion forward by posing her own question, "is this who you think will win?" (turn 15). The teacher is also an active participant generating student talk not only through some authentic questions, but also through her use of uptake questions in turns 6 and 14 ("tell me more"). While there were some opportunities for students to engage in text-based discussions like this, these were few and far between. More

frequent were indicators of dialogic instruction, like uptake questions and high-level evaluations, that happened within patterns of talk like whole class question and answer episodes.